Forum for Public Health in South Eastern Europe

Health Investigation: Analysis - Planning - Evaluation

A Handbook for Teachers, Researchers and Health Professionals (2nd edition)
Volume II

Editors:
Genc Burazeri and Lijana Zaletel Kragelj
Assistant Editor: Kreshnik Petrela and Herion Muja

Jacobs Verlag

Forum for Public Health in South Eastern Europe
A Handbook for Teachers, Researchers and Health Professionals (2nd edition)
Volume II
Health Investigation:
Analysis - Planning - Evaluation
Editors: Genc Burazeri and Lijana Zaletel Kragelj
Assistant editor: Kreshnik Petrela and Herion Muja

Health Investigation: Analysis - Planning - Evaluation

Forum for Public Health
in South Eastern Europe
A Handbook for Teachers, Researchers and Health Professionals (2nd edition)
Volume II

Editors:
Genc Burazeri and Lijana Zaletel Kragelj
Assistant editor: Kreshnik Petrela and Herion Muja

Jacobs Verlag

Project Coordinators:

Ulrich Laaser and Luka Kovacic

Editors of the 1st Edition:

Volume 1 on Health Systems and their Evidence Based Development: Vesna Bjegovic-Mikanovic and Doncho Donev

Volume 2 on Health Determinants in the Scope of New Public Health: Lidia Georgieva and Genc Burazeri

Volume 3 on Public Health Strategies: A Tool for Regional Development: Silvia Gabriela Scintee and Adriana Galan

Volume 4 on Health Promotion and Disease Prevention: Doncho Donev, Gordana Pavlekovic, and Lijana Zaletel Kragelj

Volume 5 on Management in Health Care Practice: Luka Kovacic and Lijana Zaletel Kragelj

Volume 6 on Methods and Tools in Public Health: Lijana Zaletel-Kragelj and Jadranka Bozikov

Bibliographic information published by Die Deutsche Bibliothek.
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available on the Internet at http://dnb.ddb.de

This publication has been supported by the Academic Programmes for Training and Research in Public Health in South Eastern Europe (FPH-SEE).

Hellweg 72, 32791 Lage, Germany
ISBN 978-3-89918-807-3

Preface

In order to develop the training and research capabilities for public health in South Eastern Europe a project funded by the German Stability Pact started in 2000. It was meant to support the reconstruction of postgraduate public health training programs through different activities, including the development of teaching modules. Originally planned to be on an Internet platform only, the Forum for Public Health in South Eastern Europe (FPH-SEE)[1] and the MetaNET project together with Hans Jacobs Publishing Company decided to publish this training material also as hard copy volumes. The first book was published in 2004 and the sixth one in 2010, together comprising around 3500 pages. After successful and widespread use of the teaching modules of all six books between 2004 and 2011[2], the project coordinators decided - again together with Hans Jacobs Publishing Company - to publish a 2nd fully revised edition of selected modules as e-book.

The 2nd edition has been prepared for publication in two volumes under the titles **Health: systems – lifestyles – policies** (Volume I) and **Health Investigation: analysis – planning – evaluation** (Volume II). Volume II comprises the collection of 43 teaching modules, written by 49 authors from 10 countries. The teaching modules in this book cover population health, special methods and applications as well as planning and evaluation. The authors had full autonomy in the preparation of their teaching modules. They were asked to present their own teaching/training materials with the idea to be as practical and lively as possible. Having that in mind, the reader, and the user of the modules of this book may sometimes find that some areas are not covered, some are just tackled, and some are more deeply elaborated. The role of the editors was more to stimulate the authors to write and to revise modules, than to amend or edit their content.

The project coordinators and the editors of the 2nd edition are very grateful for the continuing interest of the authors to publish their material and share their experience. We look back to more than a decade of cooperation and networking and are happy to see the fruits of this work grow ripe. We are confident that the selected 2nd edition will stabilize this success and contribute to lead South Eastern European Public Health into a future of excellence and stability.

Zagreb, 30th November, 2013
The coordinators: Professors Luka Kovacic (Croatia) and Ulrich Laaser (Germany)
The editors: Professors Genc Burazeri (Albania) and Lijana Zaletel Kragelj (Slovenia)

1 http://www.snz.unizg.hr/ph-see/index.htm

2 Zaletel-Kragelj L, Kovacic L, Bjegovic V, Bozikov J, Burazeri G, Donev D, Galan A, Georgieva L, Pavlekovic G, Scintee SG, Bardehle D, Laaser U (2012) Utilization of teaching modules published in a series of handbooks for teachers, researchers and health professionals in the frame of "Forum for Public Health in South Eastern Europe - Programmes for training and research in public health" network. Slovenian Journal of Public Health 51/4: 237-250

Editors and Project Coordinators:

Prof. Genc Burazeri MD, PhD
Faculty of Public Health
University of Medicine
Rr. "Dibres", No. 371
Tirana, Albania
E-mail: gburazeri@yahoo.com

Prof. Luka Kovacic, MD, PhD
Andrija Stampar School of Public Health
Medical School, University of Zagreb
Rockefellerova 4
10000 Zagreb, Croatia
E-mail: lkovacev@snz.hr

Kreshnik Petrela
Institute of Public Health
Rr. "Aleksander Moisiu", No. 80
Tirana, Albania
E-mail: kreshnikp@gmail.com

Prof. Dr. med. Ulrich Laaser DTM&H, MPH
Section of International Public Health
Faculty of Health Sciences
University of Bielefeld
POB 10 01 31
D-33501 Bielefeld, Germany
E-mail: ulrich.laaser@uni-bielefeld.de; laaseru@gmail.com

Prof. Lijana Zaletel-Kragelj, MD, PhD
Chair of Public Health
Faculty of Medicine, University of Ljubljana
Vrazov trg 2
1000 Ljubljana, Slovenia

Herion Muja
Institute of Public Health
Rr. "Aleksander Moisiu", No. 80
Tirana, Albania
E-mail: herionmuja@gmail.com

Table of Contents

HEALTH INVESTIGATION: ANALYSIS – PLANNING – EVALUATION

HEALTH INVESTIGATION: ANALYSIS – PLANNING – EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Introduction to the measurement of health and disease**
Module: 2.1	ECTS (suggested): 0.1
Author(s), degrees, institution(s)	Tatjana Pekemezovic, MD, PhD, Professor – Institute of Epidemiology, Faculty of Medicine, University of Belgrade, Serbia; Tatjana Gazibara, MD, Teaching Assistant – Institute of Epidemiology, Faculty of Medicine, University of Belgrade, Serbia.
Address for correspondence	Tatjana Pekmezovic Institute of Epidemiology, Faculty of Medicine, University of Belgrade, Serbia Visegradska 26A, Belgrade, Serbia E-mail: pekmezovic@sezampro.rs
Keywords	Adjustment of rates, case-fatality ratio, cumulative incidence, incidence rate, mortality rate, person-time concept, prevalence.
Learning objectives	At the end of this topic student will be able to: • describe and explain basic measures of health and disease such as incidence rate, cumulative incidence, prevalence, mortality rate, case- fatality ratio; • calculate specified rates and proportions; • understand and explain the person-time concept; • describe methods for rates adjustment and understand the principles and limitations of standardization; • understand epidemiological literature that uses and refers to the concepts outlined above.
Abstract	The epidemiological research is inquiring into the frequency of occurrence of states and events of health. The first-order focus needs to be on concepts pertaining to rates of occurrence. A distinction between prevalence (of states) and incidence (of events) is made. A population at risk must be defined clearly. Any measure of occurrence is impossible to interpret without a clear statement of the period during which the population was at risk and the cases were counted. Further on rates, one should distinguish between the overall rate and specific rates. This leads to the concept of adjustment and this, in turn, to the concept of mutually standardized rates.
Teaching methods	The teaching method recommended: • the introduction lecture relating to basic definitions and concepts; • the distribution of the literature to small group (3-4 students); • the guided discussion within each group and added explanations; • the distribution of exercises to each group; • overall discussion.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection equipment, internet connection, access to the bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written examination with calculation of rates.

INTRODUCTION TO THE MEASUREMENT OF HEALTH AND DISEASE

Tatjana Pekmezovic

Theoretical background

Introduction

The fundamental epidemiological measure is the frequency with which the events of interest (usually disease, injury, or death) occur in the population to be studied. The targets in epidemiological investigations are populations.

The frequency of event can be measured in different ways, and it can be related to different denominators, depending on the purpose of the research and availability of data (1).

Ratio, proportion, rate

Measures of health and disease are the ratios. The ratio is the value obtained by dividing one quality by another; for example, sex ratio (or male to female ratio) (2). We distinguish between proportions and rates:

1. The proportion is a type of ratio where the numerator is included in the denominator. The ratio of a part to the whole can be expressed as a (2):
   - "vulgar fraction" (1/2),
   - as a percentage (50%), and
   - as a decimal (0.5).
2. The rate is a measure of the frequency of occurrence of a disease or other health-related events. The components of a rate are:
   - the numerator (number of events),
   - the denominator (the specific period in which events occur), and usually
   - a multiplier a power of 10 (10^n).

A true rate includes the sum of time units of exposure for all people at risk (person-time concept). It is useful in small populations. In large populations, a mid-period population usually can be considered a good estimate of the average number of people at risk, for the outcome during the time period. The mid-period population, as approximation, is often used as the denominator (1). It is very important to underline that the population at risk must be defined clearly. All people who are not usually resident in that area, and those who are not at risk of the event under investigation, must be excluded from denominator (3). A difference between true rate and rate in a classical epidemiological sense is presented in a separate module in this book.

The rates usually have values less than 1, and decimals are awkward to think about and discuss. Therefore, rates are usually multiplied by a constant multiplier, either 100 or else, 1000, 10,000, 100,000, 1,000,000 in order to make the numerator larger than 1 and therefore easier to discuss (1).

Types of measures of occurrence

According to the concept of incidence and prevalence

The most frequent measures of occurrence of health-related events include incidence rate (IR), prevalence (PREV), cumulative incidence (CI), mortality rate (MR), and case-fatality ratio (CFR).

1. The incidence rate

The incidence describes the frequency of occurrence of new cases during the time period. The incidence rate (person-time incidence rate, also called incidence density) is the number of new occurrence of disease in the study population during the time period, divided by the sum of time that each person in the population remained under observation and free of disease. In other words, the denominator of incidence represents the number of people who are at risk for development of disease. The incidence rate is a direct indicator of risk of disease in a population investigated and it is a measure of efficiency of preventive measures (7). This measure is presented in detail in a separate module in this book.

2. The cumulative incidence

The cumulative incidence is the proportion of people who become diseased during a specified period of time. Both the numerator and denominator include only those individuals who at the beginning of the period are free from the disease and therefore are at risk of getting it. The cumulative incidence depends on the incidence rate and the length of the period at risk. The cumulative incidence (risk) and the incidence rate (person-time incidence rate) can be mathematically related (Equation 1):

$$CI = 1 - e^{(-I \times t)}$$ **Equation 1.**

CI = cumulative incidence
I = person-time incidence rate
t = length of follow-up

Different methods of calculation of cumulative incidence are presented in detail in a separate module in this book.
The cumulative incidence is a useful approximation of incidence rate when the rate is low, or when the study period is short (8).

3. The prevalence

The prevalence is the proportion of the population affected by a disease at a given point in time. The proportion of population that has a disease at a point in time (P) and the rate of occurrence of new disease during a period of time (I) are closely related (Equation 2):

$$P = I \times t$$ **Equation 2.**

P = point prevalence
I = incidence
t = length of duration of disease

Prevalence does not involve measurement of risk. This measure is helpful in assessing the need for health care and the planning of health services (7). In the medical and public health literature, the word prevalence is often used in two ways:

- point prevalence: implies the prevalence of a disease at a given point in time;
- period prevalence: involves the number of people who have had the disease at any time during a certain period of time.

4. The mortality rate

The mortality rate is the number of deaths in a specified period of time in a given population (a mid-period population). Mortality is a measure of risk of death in populations and efficiency of preventive measures (8). The same principles mentioned in the discussion of incidence apply to mortality: for a rate to make sense, anyone in the group represented by the denominator must have the potential to enter the group represented by the numerator.

5. The case-fatality ratio

The case-fatality ratio is the number of deaths from a disease in a specified period of time, divided by number of diagnosed cases in the same period. The case-fatality ratio is a measure of the severity of disease and efficiency of treatment procedures (9). In other words, the case-fatality ratio is the percentage of people diagnosed as having a certain disease who die within a certain time after diagnosis.

According to different types of adjustment

There are three broad categories of measures according to different types of adjustment:

- crude measures;
- specific measures, and;
- standardized measures.

1. Crude measures

The measures that apply to an entire population, without reference to any characteristics of the individuals in it are crude measures (for example, annual mortality rate from all causes of death in a country).

2. Specific measures

Specific measures may be specific according to age, sex, cause or some other characteristic (for example, annual mortality rate from breast cancer in females).

3. Standardized measures

Standardized measures are very useful when we compare two populations with different age structures. In this way, the effect of age as a confounding variable may be controlled. The essence of standardization is comparing the investigational populations with standard populations with a known age structure. The standard population is a hypothetical population, and choice of it depends on the purpose of the analysis. For international comparisons, European, or World standard populations are favoured (4).

There are two methods of adjustment, direct or indirect:

- In the **direct method of standardization**, the age-specific measures of two (or more) populations to be compared are applied to a reference population known as the standard. This is done by multiplying each age-specific measure of a population to be compared by the number of persons in the corresponding age group of the standard population. This way, one derives the expected numbers of deaths that would have occurred in populations being compared. Dividing each of the total expected numbers by the standard population leads to the adjusted, or standardized measures (5). This procedure is presented in detail in a separate module in this book.
  Direct standardization is useful to compare different areas/regions/countries with each-other and to evaluate trends over time. However, in order to be calculated, it requires age-specific rates and cannot be used in case of rare events.
- **Indirectly standardized measures** compare the actual number of events in an area with the expected number of events based on mortality measures of a standard population. This method is often used to look at differences in mortality rates, and is often referred to as standardized mortality ratio (SMR). The standardized mortality ratio is the ratio of observed to expected number of deaths, expressed as a percentage. A SMR greater than 1.00 indicates that the observed number of deaths exceeds the expected number, and a SMR less than 1.00 indicates that the observed number of deaths is less than the expected number. It can also be used to look at other events such as, for example, hospital activity. The observed figures come from the local area, and the expected numbers from applying the death rates of the standard population to the local population. The following steps are used to calculate the SMR:
  - find the age-specific death rates in the standard population;
  - find the age-specific populations in the observed area;
  - calculate the expected deaths in each of the age groups by multiplying the population in area A by the death rate in the reference population;
  - add up the number of deaths in each age group to get the total number of expected deaths.

Indirect standardisation is more robust with small numbers and avoids the distortions caused by direct standardisation based on unstable age-specific rates (3,4).

Indirect standardization is useful to determine if disease incidence/mortality is high or low in one area only and if it is a rare disease and therefore number of deaths in population groups is small. A drawback of this method is that it cannot compare SMRs with each-other unless population structures are identical, nor can it look at trends over time.

The decision to use crude, standardized, or specific measures depends on the information that an investigator is trying to obtain or impart:

- crude measures represent the actual experience of the population and provide data for the allocation of health resources and public health planning. Although they are easy to calculate and widely used for international comparisons, the fact that these values may be confounded by differences between underlying population structures make any observed differences in crude measures difficult to interpret.
- specific measures are unconfounded by that factor and provide the most detailed information about the pattern of the disease in a given population.
- standardized measures provide a summary value that removes the effect of the differences in population structures and allows for valid comparisons between groups, or with a certain group over time. The actual value of the standardized measures is meaningless, however, since it has been statistically constructed based on the choice of a standard.

Finally, depending on the nature of the information required, one or a combination of different measures can be chosen (6).

Case study

Mortality rates from breast cancer in women in two units in Belgrade

Introduction

In two urban units in Belgrade, mortality rates from breast cancer (BC) in women were as presented in Tables 1 and 2:

Table 1. Mortality rates (Mt) from breast cancer (BC) in women from urban unit A in Belgrade

Age group	No. of women	No. of deaths from BC	Mt/100,000
0-19	25,138	0	0
20-29	14,961	1	6.68
30-39	18,249	3	16.64
40-49	17,251	8	46.37
50-59	16,849	23	136.51
60-69	13,187	13	98.58
70+	9980	9	90.18
All ages	115,615	57	49.30

Table 2. Mortality rates (Mt) from breast cancer (BC) in women from urban unit B in Belgrade

Age group	No. of women	No. of deaths from BC	Mt/100,000
0-19	6722	0	0
20-29	3545	0	0
30-39	5832	1	17.15
40-49	5173	3	57.99
50-59	4770	5	104.82
60-69	6485	7	107.94
70+	5554	9	162.04
All ages	38,081	25	65.65

Comparison of the overall and age-specific mortality rates

First, we will show how overall and age-specific mortality rates from BC could be computed and compared.

The following questions could be posed:

1. Are mortality rates higher in unit A, or unit B?
2. Are there any reasons for this situation?
3. How can the difference between age-specific and crude mortality rates be explained?
4. How can the problem of comparability be overcome?
5. What is essential in standardization?
6. How can the standard population be chosen?

In the example mentioned above, we chose World population as a standard. Calculation of standardized mortality rates for BC in units A and B are summarized in Table 3:

Table 3. Calculation of standardized mortality rates for breast cancer in urban units A and B in Belgrade

		Unit A		Unit B	
(1)	(2)	(3)	(2 × 3)	(5)	(2 × 5)
Age group	Standard population	Mt/100,000	No. of expected deaths	Mt/100,000	No. of expected deaths
0-19	40,000	0	0	0	0
20-29	16,000	6.68	1.07	0	0
30-39	12,000	16.64	1.97	17.15	2.06
40-49	12,000	46.37	5.56	57.99	6.96
50-59	9,000	136.51	12.29	104.82	9.43
60-69	7,000	98.58	6.90	107.94	7.56
70+	4,000	90.18	3.61	162.04	6.48
No. of all expected deaths			31.40		32.49

Computation of standardized mortality ratios (SMR)

In continuation, we can pose a question: what are the standardized mortality rates from BC in units A and B?

For answering this question, we will use data from the following table (Table 4) (Adapted according to Hennekens & Buring, 1987) (6):

Table 4. Computation of standardized mortality ratios (Adapted according to Hennekens &Buring, 1987) (6)

Age group	Population	Mt/100,000	No. of expected deaths	No. of observed deaths
(1)	(2)	(3)	(2 × 3)	(4)
10-19	74,598	12.26	9.14	10
20-29	85,077	16.12	13.71	20
30-39	80,845	21.54	17.41	22
40-49	148,870	33.96	50.55	98
50-59	102,649	56.82	58.32	174
60-69	42,494	75.23	31.96	112
Total	534,533		181.09	436

Final calculation is presented in Equation 3.

$$SMR_{10-69years} = \frac{436}{181.09} \times 100 = 241$$ **Equation 3.**

Exercises

Teaching methods for this topic, among others, would include the distribution of different exercises in small groups of students and calculation and explanation of different measures of health-related events (all tasks are adapted from ref. 9).

Task 1: In 1997, there were 39 cases of myocardial infarction in town A among people aged 50-54 years. The number of person-time was 515,212 in that age group. Calculate the incidence rate of myocardial infarction.

Task 2: A sample including 2368 women at the age group 70-74 years was selected from the population of town B. After examination, 80 were assigned the diagnosis of rheumatoid arthritis. Calculate the prevalence of this disease.

Task 3: Of 229,400 children born in a given region, 411 had one congenital malformation at birth. Which measure of occurrence of congenital malformation can be calculated? Calculate this.

Task 4: Assume that in a population of 100,000 persons, 20 people contracted tetanus. In one year, 18 people died from that disease. Calculate the incidence rate, mortality rate and case-fatality ratio. Explain why the same disease has low incidence and mortality rates, but high case-fatality ratio?

Task 5: One research evaluated the frequency of depression among patients with dementia. In 2002, a total of 201 patients with dementia were included, of which 91 had already had depression. In the following year, 7 patients presented with symptoms of depression. Calculate incidence rate of depression in 2003.

Task 6: In a study in the country A, the frequency of stroke was measured in 228,525 women who were 30-45 years of age and free from coronary heart disease, stroke and cancer in 1997. A total of 546 stroke cases were identified during 10 years of follow-up. Calculate the cumulative incidence.

Assessment of students (types of questions):

1. The risk of acquiring a certain disease is best expressed through:
   a) prevalence
   b) incidence rate
   c) mortality rate
   d) survival rate
   e) none of the above

2. Which of the following are considered measures of the occurrence of disease:
   a) incidence rate
   b) prevalence
   c) cumulative incidence
   d) mortality
   e) all of the above

3. The incidence rate of a disease is 5 times greater in women than in men, but the prevalence shows no sex-difference. The best explanation is:
   a) The crude all-cause mortality rate is greater in women.
   b) The case-fatality ratio for this disease is greater for women.

c) The case-fatality ratio for this disease is lower for women.
d) Risk factors for developing the disease are more common in women

4. Which of the following is a good measure of the severity of an acute disease:
   a) cause-specific death rate
   b) survival rate
   c) case-fatality ratio
   d) standardized mortality rate
   e) none of the above

5. The most appropriate measures to compare disease occurrence between developed and developing countries in all age groups are:
   a) crude rates
   b) age-specific rates
   c) standardized rates
   d) survival rates
   e) all of the above

6. Age-adjusted death rates are used to:
   a) correct death rates for errors in the statement of age
   b) determine the actual number of deaths that have occurred in specified age groups in a population.
   c) correct death rates for missing age information
   d) compare deaths in person of the same age group
   e) eliminate the effects of difference in the age distributions of populations in comparing death rates

7. Direct standardization is used when investigating:
   a) rare diseases
   b) disease occurrence in populations without age-specific rates
   c) disease rates in one region only
   d) disease occurrence in populations with age-specific rates
   e) populations of the same age structures

8. Indirect standardization method (circle the answer which is NOT correct):
   a) is expressed as percentage
   b) does not require age-specific rates of the population
   c) is used for comparisons in the same age structures
   d) is used for comparisons of disease rates in one region only
   e) cannot be applied for time trends

References

1. Jekel JF, Elmore JG, Katz DL. Epidemiology, Biostatistics and Preventive Medicine. Philadelphia: W.B. Saunders Company, 1996.
2. Last JM, editor. A dictionary of Epidemiology. New York, Oxford, Toronto: Oxford University Press, 2001.
3. Dos Santos Silva I. Cancer epidemiology. Principles and Methods. Lyon: IARC, 1999.
4. Jensen OM, Parkin DM, Mc Lennan R, Muir CS, Skeet RG. Cancer registration. Principles and Methods. Lyon: IARC, 1991.
5. Radovanovic Z, Jevremovic I, editors. Epidemiology. Practical sessions. Belgrade: School of Medicine, University of Belgrade, 2000.
6. Hennekens CH, Buring JE. Epidemiology in Medicine. Boston/Toronto: Little, Brown and Company, 1987.
7. Mausner JS, Kramer S. Epidemiology - an introductory text. Philadelphia: W.B. Sounders Company, 1985.
8. Coggon D, Rose G. Barker DJP. Epidemiology for the uninitiated. London: BMJ Publishing Group, 1997.
9. Gordis L. Epidemiology. 4th edition. Philadelphia: W.B. Sounders Company, 2009.
10. Ahlbom A. Introduction to Modern Epidemiology. Stockholm: Epid. Res. Intern, 1984.

Recommended readings

1. Last JM, editor. A dictionary of Epidemiology. New York, Oxford, Toronto: Oxford University Press, 2001.
2. Beaglehole R, Bonita R, Kjellstrom. Basic Epidemiology. Geneva: WHO, 2007.
3. Fleiss K. Statistical methods for proportions and rates. New York: John Willey and Sons, 1981.
4. Rothman KJ, Greenland S. Modern Epidemiology. Philadelphia: Lippincott Raven Publishers, 1998.
5. Jekel JF, Elmore JG, Katz DL. Epidemiology, Biostatistics and Preventive Medicine. Philadelphia: W.B. Saunders Company, 1996.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Organizing and describing data**
Module: 2.2	ECTS (suggested): 0.4
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Lijana Zaletel-Kragelj Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Data description, data organization, ratio, statistical distribution, typical value.
Learning objectives	After completing this module students should: • know how to organize data for statistical and epidemiologic description; • be familiar with basic statistical description of data (frequency distribution, typical values of distribution); • be familiar with basic epidemiologic description of data, and • be aware of existence of different ratios, used in epidemiology.
Abstract	As in any other profession, in public health (PH) too, the research process (in this profession the research issues are different kinds of health problems of a population and their determinants) takes a very important role. Organizing and describing data is the very beginning of this process. The module is describing basic principles of statistical and epidemiologic description of the data.
Teaching methods	An introductory lecture gives the students first insight in characteristics of organization and description of data, statistics and epidemiology. The theoretical knowledge is illustrated by three case studies. After introductory lectures, students first read carefully the theoretical background of this module and complement their knowledge with recommended readings. Afterwards, they work in pairs with data sets and perform two extensive tasks. Students use computer programmes to complete their exercise. Students are stimulated to compare their results with results of the other pairs and discuss the possible differences.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer for 2-3 students), LCD projection, access to the Internet and statistical programmes (recommended: SPSS); • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on analysis of a given data set.

ORGANIZING AND DESCRIBING DATA

Lijana Zaletel-Kragelj

Theoretical background

Introduction

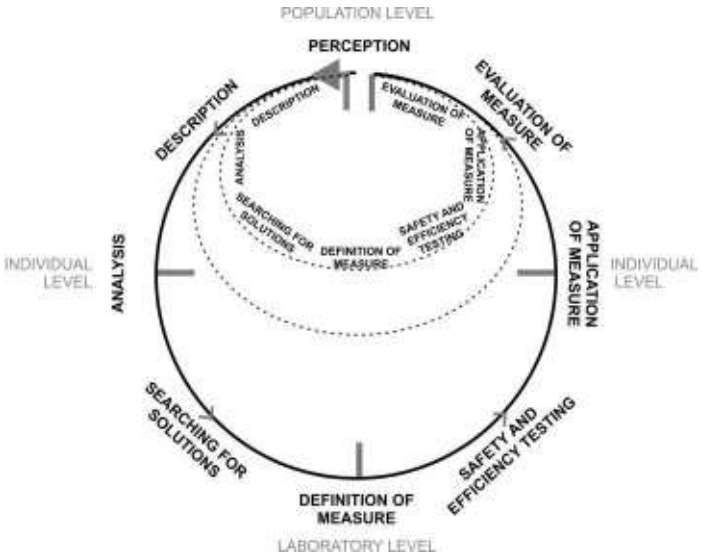
Why to organize and describe the data in public health

As in every other profession, in public health (PH) too, the research process takes a very important role. In PH the research issue is different kinds of health states/problems of a population (i.e. diseases, disabilities, injuries, deaths), and their determinants (1-6). Organizing and describing data is the very beginning of this process.

Research process in public health

The phases of this process are similar to other research processes in medicine (i.e. in clinical medicine, laboratory medicine etc.) and in fact represent a very important part of the whole process of solving health problems. The phases are as follows (Figure 1):

Figure 1. The levels of research process in public health



1. perception of a health problem,
2. description of the problem,
3. analysis of the problem,
4. searching for possible measures for solving the problem,
5. the final definition of the measure,
6. testing safety and efficiency of the measure on human beings,
7. mass application of the measure on individuals,
8. observation of a long-term safety and efficiency of the measure.

In the first part of the process we are describing and analyzing the problem from the population level through level of an individual to the laboratory level (Figure 1) aiming at discovering the most appropriate measure for solving it. In the second part, first testing of safety and efficiency of the measure at the level of an individual before mass application takes its role, and afterwards the evaluation of efficiency at the population level. In solving some health problems the individual level could be skipped (in phenomena which could not be measured at an individual level like different kinds of environmental or community phenomena).

In the process of organizing and describing data in PH research, the statistical methods take a very important role. The relationship between PH, epidemiology and statistics is as follows:

1. PH is defined as one of the organized efforts of a society to protect, promote, and restore the people's health. It is the combination of sciences, skills, and beliefs that is directed to the maintenance and improvement of the health of all the people through collective or social actions (1,4,7). One of these sciences, also being one of the important branches of medicine itself, is epidemiology.

2. Epidemiology in its broadest sense is defined as the study of the distribution of health states (different kinds of diseases or other phenomena related to the health of the people) and their determinants in specified populations, and its application to the control of health problems (1,4,7). Statistical methods represent one of the most powerful tools in epidemiology.
3. Statistics is defined as the science and art of collecting, summarizing and analyzing data that are subject to random variation (1). It is represented by a huge set of different methods adequate for different situations. Statistical methods take their role in (1,4):
   - description of health phenomena - descriptive statistics - and are used in descriptive epidemiology (activities to study occurrence of disease or other health-related characteristics in human populations; it is concerned with where, when and how frequent such phenomena are), and
   - analyzing of health phenomena - the methods of analytical statistics - and are used in analytic epidemiology (usually concerned with identifying or measuring the effects of risk factors, or with the health effects of specific exposures).

Organizing data

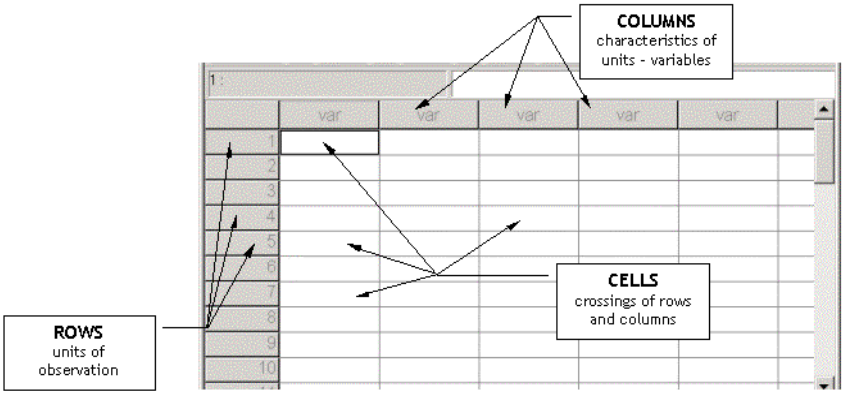
Data matrix

For good-quality research, it is of basic importance to have data well-organized and prepared for both description and analysis. As the methods for both kinds of activities are statistical methods, it is very important to follow the rules of preparing the data in an adequate structure for statistical analysis.

The appropriate structure is a data matrix (1). This is the structure in which data of all observational units and all observed attributes of units are organized in a table (Figure 2). The basic element of this table is a cell. The cells are organized in a matrix with rows and columns. The meaning of elements of this table is as follows:

1. Cell – the record of a piece of information (lat. datum) on single attribute (variable) of a single unit of observation (statistical unit),
2. Row - the record of values of all variables for a single unit,
3. Column - the record of values of all units for a single variable.

Figure 2. Organization of data for statistical description and analysis



Statistical description of data

Overview of foundations of statistics

Basic statistical concepts

There exist four basic concepts in statistics (1,4,8,9):

1. Statistical population – the whole collection of units of phenomenon under study subjected to statistical methods,
2. Statistical unit – every single element (member) of statistical population,
3. Statistical variable – every single characteristic (attribute, phenomenon) of statistical unit under study,
4. Statistical sample – a selected subset of a statistical population; it may be random, or non-random; it may be representative, or non-representative.

All these concepts are closely related to each-other:

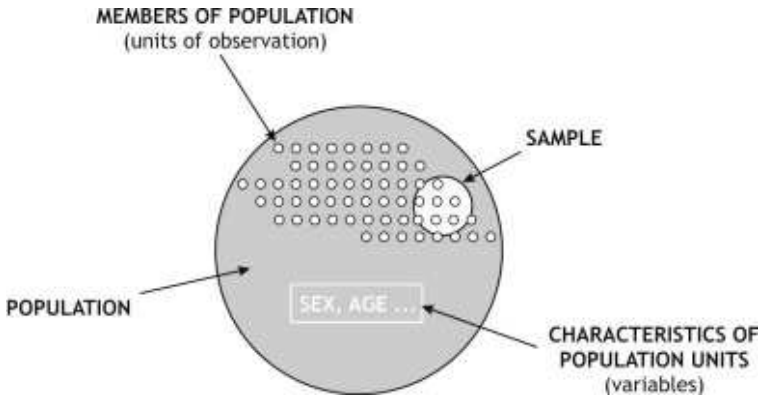
- when we are performing statistical observation of a certain phenomenon, the subject of interest is a whole mass of members, called statistical population,
- one single member of this mass is called unit of observation or statistical unit,

- the units have the attributes of their own. As these attributes can have different values (they vary), we call them variables,
- usually we cannot observe the whole population under study, so we draw a sample from the population. In that case, we describe first the statistical features of the sample and then we generalize to the overall population.

The relationship among these basic concepts is also shown in Figure 3.

The key concept in statistics is the concept of statistical variable or more precisely the concept of random variable, or variate (1,10,11). According to Last et al., a variate is a variable that may assume any of a set of values, each with a pre-assigned probability (1).

Figure 3. Basic concepts in statistics and their respective relationships



Concepts related to statistical activities

In statistics, we can perform the following kinds of activities (8,9):

1. Statistical description – the process of summarizing the characteristics of data under study (at the sample or population level); we call this process descriptive statistics,
2. Statistical relationship analysis - the process of analysis of relationships between dependent (effect) and one or several independent (causes) variables (phenomena),
3. Statistical inference – the process of generalization from sample data to population, when the observation is not performed in a total population, but only in a (representative) sample, usually with calculated degrees of uncertainty; we call this process inferential statistics.

If we observe the total population, we perform only the methods of descriptive statistics. When only the sample is available we usually need to perform description and inference, whereas relationship analysis could be performed in both situations.

Which methods are to be used depends on statistical features of variables under research.

Concepts related to statistical variables

Statistical description and inference are closely related to the concept of statistical variable. Here we shall introduce some other concepts, also closely related to it.

Values of variables and their distribution

The first two important concepts are:

1. variable values – every single variable can take two or more different values,
2. distribution of variable values – the complete summary of the frequencies of the values of a single variable (some of the values are more frequent than the others); it can tell the number or the proportion of the whole group of observations to be of each value out of all observations.

Classifying variables

The variables, or more precisely their values, could have various statistical features. Regarding these features they could be classified in several ways (1,4,12):

1. Regarding the expression of their values to:
   - numerical variables – variables, values of which are expressed by numbers (e.g. weight, number of patients per day),
   - categorical (qualitative, attributable) variables or attributes – variables, values of which are expressed only by description (e.g. sex),
2. Regarding the possibility of infinitive number of their values to:

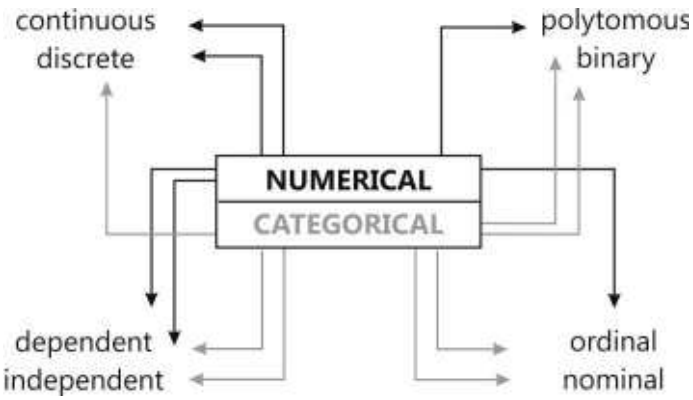
- continuous variables – variables with potentially infinite number of possible values along a continuum (e.g. weight, height),
- discrete variables – variables values of which could be arranged into naturally or arbitrarily selected groups of values (e.g. number of patients per day),

3. Regarding the ordinality of values to:
   - ordinal variables – variables values of which are classified into ordered categories (e.g. social class),
   - nominal variables - variables values of which are classified into unordered categories only by equality or inequality (e.g. race, religion, country),
4. Regarding the number of distinct values to:
   - dichotomous or binary variables – variables with only two possible values, often contain information of having the characteristic of interest or not,
   - polytomous variables – variables with more than two possible values,
5. Regarding the interrelationship between two or more variables to:
   - dependent variables – variables values of which are depending on the effect of other variables (independent variables) in the relationship under study,
   - independent variables – variables that are hypothesized to influence the values of other variables (dependent variables) under study.

All these classifications could be related to each-other. When we put the classification on numerical and categorical variables in the central position and link it to all other classifications, then we get (Figure 4):

- numerical variables are continuous or discrete, only ordinal and polytomous and they could be dependent or independent,
- categorical variables are only discrete, dichotomous or polytomous, ordinal or nominal and they could be dependent or independent.

Figure 4. Various classifications of variables and the linkage of classification into numerical and attributable variables with all other classifications



This linkage leads to classification of types of variables.

Types of variables

Usually we are classifying variables into four main categories (5,13):

1. Numerical continuous variables,
2. Numerical discrete variables,
3. Categorical ordinal variables, and
4. Categorical nominal variables.

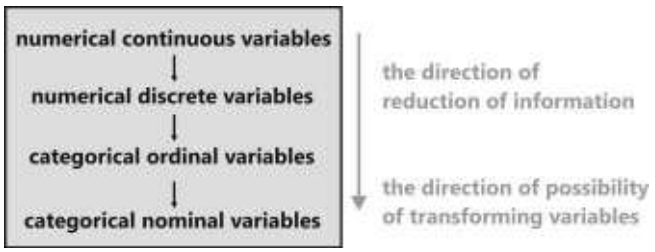
This sequence of types of variables represents also their hierarchy regarding the amount of information encompassed in each of them. In the direction from numerical continuous to categorical nominal variables, the amount of information is decreasing. In this direction also the transformation from one type to another is possible but it is not possible in the opposite way (Figure 5).

What does the amount of information means and what means reducing it, would be more understandable in an example presented in "Case study 2".

In the phase of planning and designing the study it is very important to be aware that the information could be reduced anytime but could never become more precise unless we acquire it once again.

Classifying the variables into right types is very important for deciding which method is to be used in describing variables, and examination of the relationships between them.

Figure 5. Hierarchy of types of variables regarding the amount of information encompassed in each of them



Concepts related to probability distributions

Some of variable values are more frequent than the others – they are more probable. The way how frequent the values of the particular variable are is called probability law of the variable. The distributions of values of variables are therefore called probability distributions. According to Last et al., probability distribution for a discrete random variable is the function that gives the probabilities that the variable equals each of a sequence of possible values, while for a continuous random variable is often used synonymously with the probability density function – the frequency distribution of a continuous random variable (1). We could roughly classify probability distributions into two groups:

1. Empirical probability distributions – distributions observed in real situation,
2. Theoretical (mathematical) probability distributions – mathematical idealization of distributions observed in real situations.

By far, the most important theoretical probability distribution is known as Normal or Gaussian distribution (1,4,8,9,11,14).

Other also important theoretical distributions (all of them are families of similar distributions, varying with regard to the number of observations) are (1,4,8,9):

- Student's t distribution,
- binomial distribution,
- chi-square distribution,
- Poisson distribution,
- Fisher's F distribution.

Theoretical probability distributions are very important as many statistical methods are based on the assumption that the observed data are a sample drawn from a population with known distribution. If such assumption is reasonable (it could never be checked whether it is true) the use of statistical methods becomes simple. But here we have to warn that if the assumption of distribution under study is not reasonable and we proceed with the activities, we could make the misleading conclusions.

Process of statistical description of data

Statistical description of data is a set of consecutive procedures used for describing the empirical distributions in an agreed way. The result of these procedures is:

1. Description of a shape of distribution, and
2. Determination of measures which summarize the features of the shape.

When statistically describing data we can choose between methods that make assumptions on theoretical probability distributions, called parametric methods (the origin of this term will be discussed later) and those which make no such assumptions, called non-parametric or distribution-free methods (8,9,14).

Presenting Data

The presentation could be numerical or graphical:

1. Numerical data presentation - ordered series and frequency distribution.

The very first step in describing data statistically is to put data in order by making first an array and then a frequency distribution table:

- ordered series or an array (10) - arrangement of values of a variable in order, usually from the lowest to the highest value,
- frequency distribution - we summarize the frequency of every single value of a variable in ordered series in a table in which we usually insert two kind of frequencies:
  - absolute frequency called usually simple frequency - the number of units with particular value of a variable,
  - relative frequency - ratio between the number of units with particular value of a variable and all units under study; it could be expressed as a proportion (decimal fraction) or as a percentage (different ratios are discussed later in this chapter),

– in some statistical programs also the third kind of frequency could be found, called cumulative frequency. This is the number of units with values less than or equal to each value. It could be expressed also as a relative measure - relative cumulative frequency.

The example of a frequency distribution is given in "Case study 2".

2. Graphical data presentation

The frequency distribution table could be useful for determination of some of data distribution features like the lowest and the highest value and thus also of the range of values, but not for all of them. The graphical data presentation is thus obligatory.

Basic graphical presentation of data in statistics is a chart of bars organized in such a way that values (categories) of a variable are listed along the x axis of the chart and their frequencies (absolute or relative) along the y axis. The area of every single bar is proportional to the frequency of the value it represents. It could be divided in two main forms (1,4,9,11):

- ordinary bar chart – the bars are lying separately; it is mostly used for presentation of attributable data,
- histogram – the bars are connected one to another; it is used for presentation of frequency distributions of numerical data; if there are many different values (continuous data) it is desirable to group observations before constructing a histogram in order to get a better visual impression of the observed distribution. An example of histogram is given in Case study 2.
- if we connect the centres of bars of a histogram at their upper part, we get the polygon called frequency polygon. When the bars are very numerous and very narrow (continuous data arranged in very small intervals) we can smooth the polygon. So we get the curve called probability density curve.

For understanding of the principles of statistical methods, it is the most important: when the relative frequency (a proportion or percentage) is used in graphical presentation of a distribution the sum of areas of all bars equals to 1 or 100%, regardless the type of bar chart is used – the ordinary one or histogram. Also the entire area under every probability density curve equals to 1 or 100%.

Data presentation by graph shows us clearly the shape of the distribution under study. This step of data presentation is very important for deciding which statistical methods are to be used for statistical description or/and inference in numerical variables.

Describing a distribution

When the graphical presentation of the shape of a distribution is done it should be described. The shape itself depends on a number and features of the place of highest density (peak). We say that distributions have diverse statistical features. Regarding these features they could be classified in several ways (4,8,9,11,14):

1. Regarding the number of peaks to:
   - unimodal – distributions with a single peak,
   - bimodal – distributions with two peaks,
   - polymodal – distributions with more than two peaks,
2. Regarding the shape of the peak to:
   - bell shaped – distributions in which extreme values tend to be less likely than values in the middle of the ordered series,
   - uniform – distributions in which all values have the same frequency,
3. Regarding the symmetry to:
   - symmetrical,
   - asymmetrical,
4. Regarding the inclination of the peak or skewness (when the distribution is not symmetrical) to distributions with:
   - positive skewness – distributions with an extended right hand tail (lower values more likely),
   - negative skewness – distributions with an extended left hand tail (higher values more likely),
5. Regarding the flatness or peakedness in symmetrical distributions to:
   - platykurtic – distributions with more flat peak than in normal distribution,
   - mesokurtic – distributions with the similar flatness of peak as in normal distribution,
   - leptokurtic – distributions with higher and slimmer peak than in normal distribution.

Usually we are the most interested in first four features.

By representing the distribution graphically we would like to get the impression if the empirical distribution under study is similar the normal distribution which is unimodal, bell shaped and symmetric. If it is case, then in determination of the measures which summarize the features of the shape parametric methods for statistical description will be used, otherwise the non-parametric ones will be used.

Summarizing the distribution features

When describing the distribution of a numerical variable, continuous or discrete, we summarize its features also by special measures called typical values or measures of location of a distribution or shortly measures of location (4,8,9,11,14).

1. Types of typical values

The most well known typical values are the following ones (4,5,8,9,11,14):

- measures of central tendency – the term includes several characteristics of the distribution of sets of values at or near the middle of the set; the principal measures of central tendency are:
- mean (average) – the sum of values of a variable for each observation, divided by the number of observations,
- median – a point in the ordered series which divides it into two parts of equal number of units, half of them falling below and half above this point,
- mode – the most frequent value in the set of observations,
- measures of dispersion or variation or spread of units around the centre of the distribution:
- minimum and maximum - the lowest and the highest value of a distribution,
- range – the difference between the minimum and the maximum,
- variance – sum of the squares of the deviations from the mean, in population divided by the number of observations,
- standard deviation – positive square root of the variance.
- subgroups, based on an array, with equal number of units; in any case the number of quantiles is one less than the number of corresponding equal parts; centiles are dividing the ordered series to hundredths and there are ninety-nine of them, deciles are dividing it to tenths and there are nine of them, quartiles are dividing it to quarters and there are three of them (median is also a quantile, dividing an ordered series to halves); for describing the spread we usually use quartiles (1^{st} and 3^{rd}) or certain centiles (25^{th} and 75^{th}),

We could classify typical values also in parametric and non-parametric ones:

- parametric typical values – (measures that are basing on normal distribution) mean as a measure of central tendency and variance and standard deviation as measures of dispersion are called parametric measures,
- non-parametric typical values – not basing on theoretical distributions.

Which set of typical values is the most appropriate for certain distribution is to be decided after observing the shape of the distribution shown by the histogram. The decision should be made not only on the shape of the distribution but also on the number of observations and whether the inferential methods would be performed. The summary about possible decision in some typical situations is shown in Table 1. An example of presentation of typical values is presented in Case study 2.

Table 1. Which typical values could be chosen in some typical examples of distributions

Shape of distribution	Other important characteristics	Typical values	
		Measure of central tendency	Measure of dispersion
Symmetrical or almost symmetrical Bell shaped		Mean	Standard deviation Minimum and maximum
Slightly asymmetrical Bell shaped	Large number of units	Mean	Standard deviation Minimum and maximum
	Small number of units	Median	Quartiles Minimum and maximum
Strongly asymmetrical	Only description	Mode	Minimum and maximum
	Inference planned	Median	Quartiles Minimum and maximum

2. Typical values in populations and samples

We can perform statistical description in populations as well as in samples. So we determine the typical values as the summary measures at both levels. Here we have to emphasize that typical values at the sample level are not the same as the typical values at the population level. In fact, in a process of inferential statistics we infer from sample characteristics to population characteristics from which the sample was drawn, that means that we infer from the values of statistics to the values of parameters. To distinguish these measures between both levels we have different names for them, and also the labeling is different. Some selected representatives of typical values are (Figure 5):

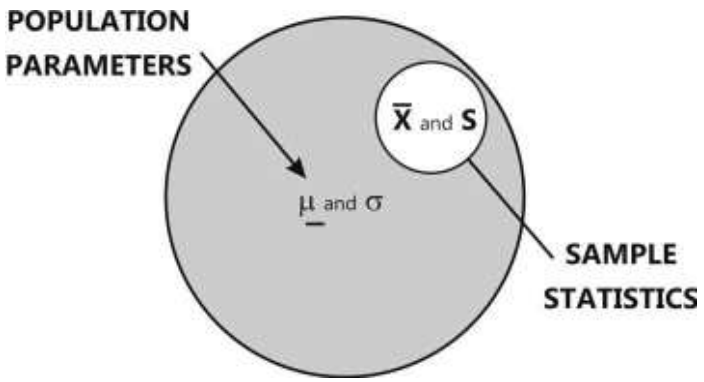
- statistics - typical values in samples:

- mean (average) of a sample, usually labeled as " $\bar{X}$ ",
- standard deviation of a sample, usually labeled as "s",
- proportion of a sample, usually labeled as "p",

• parameters - typical values in populations:

- mean (average) of a population,
- standard deviation of a population,
- proportion of a population.

Figure 5. Labeling of some typical values in populations and samples



Strict distinguishing between statistics and parameters is a basis for understanding the methods of statistical inference.

Epidemiologic description of data

Mathematical foundations of epidemiologic measurement

Basic tool for any kind of epidemiologic observation or research is quantification of frequency of health phenomena. In principle it is very similar to statistical process, but in epidemiologic measurement the emphasis is on discrete type of data, usually binary (e.g. disease is present or not, people are exposed to the certain risk factor or not) (3-5,7,15).

The frequency of a binary event could be expressed as:

- an absolute frequency or
- a relative frequency.

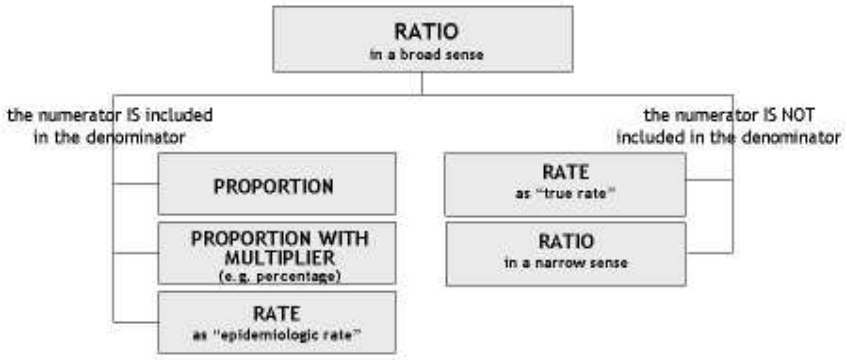
In public health, both, absolute and relative frequency measures convey important information, although relative measures seem to be frequently used. The probable reason is that relative frequency measures are important in comparisons (e.g. between two or more population groups, between two or more populations etc.), while absolute measures are important in health care planning (e.g. number of hospital beds needed for treatment of certain group of health states).

Relative frequency measures are defined as a ratio between two data. In its broadest sense the ratio is a result of dividing one quantity by another (1,15,16). One quantity is representing a numerator and the other a denominator in this relationship. The term "ratio" is a general term of which rate, proportion, percentage, etc., are subsets (1,15). The numerator and denominator need not be related (5). One of the most important features is if the numerator is included in the denominator in calculation of the ratio. Regarding the relationship of the numerator and the denominator there exist different types of ratios which could be grouped in two main groups (Figure 6) (1,5,15,16):

1. Ratios in which the numerator is included in the denominator:
   - proportion,
   - proportion with multiplier (e.g. percentage in which multiplier is 100), and
   - rate in epidemiologic sense or epidemiologic rate.
2. Ratios in which the numerator is not included in the denominator:
   - rates as "true rates",
   - ratios in a narrow sense.

The difference between different ratios will be presented using the same set of data in "Case study 3".

Figure 6. Types of relative frequency measures in epidemiology



Types of relative frequency measures in epidemiology

Ratios in which the numerator is included in the denominator

1. Proportion

Proportion is the most simple relative frequency measure (17). It is the ratio of a part to the whole (1,5,17). According to Last et al., the important difference between a proportion and a ratio is that the numerator of a proportion is included in the denominator, whereas this is not necessarily so for a ratio (1). Proportion is calculated by using following equation (Equation 1):

$$\text{Proportion} = \frac{N_{events}}{N_{total}} \qquad \textbf{Equation 1.}$$

N_{events}= number of events of observed phenomenon (part of a whole)
N_{total} = number of all possible events of observed phenomenon (a whole)

A proportion could be expressed as a "vulgar fraction" (e.g. ½) or as a "decimal fraction" (e.g. 0.5) (1).

By definition, a proportion (p), if decimal, must be in the range $0 \leq p \geq 1$ (1).

A proportion is dimensionless since numerator and denominator have the same dimension, obtained through algebraic cancellation (1). If numerator and denominator are based upon counts (e.g. in our dataset), the originals are also dimensionless. Calculating of this kind of relative frequency measure is presented in Case study 3.

2. Proportion with a multiplier

A proportion could be multiplied by a factor K (1,5,17). A multiplier is usually a power of 10 (100, 1.000, 10.000…). Its role is mainly to convert the decimal fraction to a whole number. It is calculated by using following equation (Equation 2):

$$\text{Proportion}_{\text{with a multiplier}} = \frac{N_{events}}{N_{total}} \times K \qquad \textbf{Equation 2.}$$

N_{events}= number of events of observed phenomenon (part of a whole)
N_{total} = number of all possible events of observed phenomenon (a whole)
K = multiplier (100, 1000, 10.000, 100.000…)

Which multiplier is to be used depends on a given situation (e.g. in World Health Organization Health for All Database multiplier 100.000 is mostly used).

Typical representative of this kind of relative frequency measures is a percentage, in which multiplier is 100 (17) (Equation 3):

$$\text{Percentage} = \frac{N_{events}}{N_{total}} \times 100 \qquad \textbf{Equation 3.}$$

N_{events}= number of events of observed phenomenon (part of a whole)
N_{total} = number of all possible events of observed phenomenon (a whole)

Calculation of this kind of relative frequency measure is presented in Case study 3.

3. Rate in a classic epidemiologic sense

Before trying to explain what means the term "rate in a classic epidemiologic sense" we need to discuss the term "rate" itself. To a non-epidemiologist, rate means how fast something is happening or going, for example, the speedometer of a car indicates the car's speed or rate of travel in miles or kilometres per hour (5). This rate is always reported per some unit of time. Consecutively, some epidemiologists restrict use of the term "rate" to similar measures that are expressed per unit of time. For these epidemiologists, a term "rate" describes how quickly disease

occurs in a population. These measures convey a sense of the speed with which disease occurs in a population (5). But this kind of ratio is a ratio in which the numerator is not included in the denominator and it will be discussed later.

Other epidemiologists use the term "rate" more loosely, referring to proportions with case counts in the numerator and size of population in the denominator as rates (5). For this kind of ratios we are using the term "rate in a classic epidemiologic sense". If these rates are referring to a specified period of time, they are calculated as a proportion with multiplier and specified period of time as a compulsory element by using Equation 4. An example of this kind of rate is an incidence rate in a classic epidemiologic sense. This measure will be in details discussed in a separate module of this book.

$$\text{Rate}_{\text{epidemiologic}} = \frac{N_{\text{events (in a specified time period)}}}{N_{\text{total}}} \times K \qquad \textbf{Equation 4.}$$

N_{events}= number of events of observed phenomenon (part of a whole) in a specified time period
N_{total} = number of all possible events of observed phenomenon (a whole) – at risk for occurrence of the event at the beginning of a specified time period
K = multiplier (100, 1000, 10.000, 100.000…)

In this loose usage the time component is not always referring to a period of time in which the outbreak of new cases of health phenomenon under observation is followed-up. Sometimes is referring to a number of cases in a specific point in time (Equation 5):

$$\text{Rate}_{\text{epidemiologic}} = \frac{N_{\text{cases (in a specified point of time)}}}{N_{\text{total}}} \times K \qquad \textbf{Equation 5.}$$

N_{cases}= number of events of observed phenomenon (part of a whole) in a specified point of time
N_{total} = number of all possible events of observed phenomenon (a whole) in a specified point of time
K = multiplier (100, 1000, 10.000, 100.000…)

An example of this kind of rate is a prevalence rate. This measure will also be discussed in details in a separate module of this book. Calculation of this kind of relative frequency measure is presented in Case study 3.

In epidemiology, and especially in vital statistics, this kind of measures is essential for comparing health phenomena between different populations (1,12).

Ratios in which the numerator is not included in the denominator

In ratios in which the numerator is not included in the denominator could be of two kinds. First are those in which the numerator and denominator are completely different variables (5). For this kind of ratios we use the term "true rate". In others, the numerator and denominator are different categories of the same variable. We use the term "ratio in a narrow sense" for this kind of ratio.

1. True rate

This kind of rates refers to ratios representing changes in two quantities, where the two are separate and distinct quantities. In its precise usage a rate is the ratio of a change in one quantity to a change in another quantity, with the denominator quantity often being time (18,19). A classic example of a rate is velocity, which is a change in location divided by a change in time. Dimensionality of this kind of ratio is obtained through combination of dimensions of the numerator and the denominator (e.g. km/h). In epidemiology a representative of this kind of ratio is for example so called incidence rate as a true rate. The detailed description if this measure is out of the scope of this module. It will be presented in a separate module of this book.

2. Ratio in a narrow sense

There exist also ratios that could not be classified in none of the previously presented ratios. For example, in epidemiology ratio in which the numerator and denominator are different categories of the same variable is rather frequent kind of measure. It could be simply the ratio between males and females, or persons 20-29 years and 30-39 years of age (5). The other example is a ratio in which we are relating events of an observed health phenomenon to non-events (ratio between the number of people with observed phenomenon and the number of people without it) (Equation 6). Calculation of this kind of relative frequency measure is presented in Case study 3.

$$\text{Ratio} = \frac{N_{\text{events}}}{N_{\text{non-events}}} \qquad \textbf{Equation 6.}$$

N_{events}= number of events of observed phenomenon (part of a whole)
$N_{non\text{-}events}$= number of non-events of observed phenomenon (part of a whole)

Important considerations in epidemiologic research

When we are observing the frequency of specified health phenomenon, we have first precisely to define:

1. If the study is cross-sectional or longitudinal.
   Cross-sectional study examines the phenomena in a point of time or very short period of time (e.g. a couple of weeks) while longitudinal examines it over a long period. In cross-sectional studies we are usually studying the frequency of all cases of observed phenomenon, while in longitudinal the frequency of new cases (1,4,5,7,15,17).
2. Which quantity represents the numerator and which the denominator in the equation and if the numerator is included in denominator.
3. What is the unit under observation. In epidemiology it is not necessary that the unit of observation is a person, it could be for example an episode of a health state. One should be aware in interpretation. Frequently health indicators are measuring health care services load (that is dependent also on health care services availability and accessibility, and health care services use demands of the population) and not a burden of disease in the population

Some important epidemiologic concepts

Outcome and exposure

First two important concepts are the concept of "outcome", and the concept of "exposure".

1. Outcome

Outcome is any possible disease or other health phenomenon or event related to health. It is a result of influence of an exposure to another phenomenon (1,4,5,7).

2. Exposure

Exposure is a process by which an agent (risk factor) comes into contact with a person, and provokes the relevant outcome, such as a disease (1,4,5,7).

Case, control and cohort

Other important concepts are the concepts of "case", the concept of "control", and the concept of "cohort".

1. Case (case-patient)
   In epidemiology a case is mostly defined a person identified as having the health condition under observation (1,4,5,7).
2. Control (control person)
   Controls are a group of persons with whom comparison is made in certain types of epidemiologic studies (e.g. in "case-control" studies and "randomized clinical trials") (1).
3. Cohort

In it broad sense, the tem "cohort" describes any designated group of persons followed over a period of time, as in a "cohort study" (1). We distinguish two types of cohorts in this sense (1):

- fixed cohort – in which no additional membership is allowed after beginning of the study, and
- dynamic cohort – which gains and/or loses its members during the observation time.

In other meaning it is a part of population, born during a particular period and identified by date of birth.

Probability, risk and odds

At the end, we need to present also the concepts of probability, risk, and odds.

1. Probability

In a statistical sense probability is quantification of likelihood of an event or a quantitative description of the likely occurrence of a particular event (9,12,20,21). It is conventionally expressed on a scale from 0 to 1 (a rare event has a probability close to 0, while a very common event has a probability close to 1). The probability of an event has been defined as its long-run relative frequency, defined as a ratio between number of events and total number of all possible events (Equation 7) (1,9,20):

$$p_{event} = \frac{N_{events}}{N_{total}} \qquad \textbf{Equation 7.}$$

P_{event} = probability for occurrence of observed phenomenon
N_{events} = number of events of observed phenomenon (part of a whole)
N_{total} = number of all possible events of observed phenomenon (a whole)

In fact, it is a ratio of a type "proportion" (numerator is included in the denominator) and as such could be expressed as a vulgar fraction, a decimal fraction, or as a percentage. Relative frequency expressed as a proportion of a sample is an estimate of the probability of observed phenomenon in a population.

Calculation of probability is presented in Case study 3.

2. Risk in a statistical sense

In a statistical sense risk is probability that the expected event does not occur. It could be expressed as (Equation 8) (12):

$$R_{statistical} = p_{non-event} = 1 - p_{event}$$ **Equation 8.**

$R_{statistical}$ = risk in a statistical sense
$P_{non-event}$ = probability for non-occurrence of observed phenomenon
P_{event} = probability for occurrence of observed phenomenon

This measure could be expressed as a vulgar fraction, a decimal fraction, or a as percentage as well.
The sum of probability for expected event to occur and probability that it does not occur (risk) is 1 or 100%.

3. Risk in a classic epidemiologic sense

In an epidemiologic sense the definition of "risk" is different. It is defined as a probability for an unfavourable health outcome (e.g. disease), or some other unfavourable phenomenon related to health (e.g. smoking or other unhealthy behaviour), to occur (Equation 9) (1,9).

If we are more precise, in epidemiology, the term "risk" is generally used to mean the probability that an unfavourable event (e.g., that a person will be affected by, or die from, an illness, injury, or other health condition) will occur in a given time interval (5,18). In its epidemiologic usage, risk is a conditional probability, because it is the probability of experiencing an event or becoming a case conditional on remaining "at risk" (eligible to become a case) and "in view" (available for the event to be detected) (5,18).

$$R_{unfavourable\ health\ outcome} = p_{unfavourable\ health\ outcome}$$ **Equation 9.**

$R_{unfavourable\ health\ outcome}$ = risk for an unfavourable health outcome
$p_{unfavourable\ health\ outcome}$ = probability for an unfavourable health outcome

This measure is presented in details in a separate module in this book.

4. Odds

Odds are defined as the ratio of the probability of occurrence of an event to that of non-occurrence (or the ratio of the probability that something is so to the probability that it is not so) (1) (Equation 10).

$$O_{event} = \frac{p_{event}}{1 - p_{event}}$$ **Equation 10.**

O_{event} = odds for occurrence of observed phenomenon
P_{event} = probability for occurrence of observed phenomenon

In epidemiology, if we define the probability of occurrence of an unfavourable event as a risk, it is also defined as the ratio of the risk of occurrence of a disease to that of non-occurrence (Equation 11):

$$O_{unfavourable\ healthoutcome} = \frac{R_{unfavourable\ healthoutcome}}{1 - R_{unfavourable\ healthoutcome}}$$ **Equation 11.**

$O_{unfavourable\ health\ outcome}$ = odds for an unfavourable health outcome
$R_{unfavourable\ health\ outcome}$ = risk for an unfavourable health outcome

As in the presented equation the quantity "total number of possible events" is included in both, numerator and denominator, it could be reduced through algebraic cancellation. In this case we get a new equation (Equation 12). From this equation we can clearly see, that odds are the ratio of a type "ratio in a narrow sense" - numerator is not included in denominator, and the numerator and denominator are different categories of the same variable. Calculation of odds is presented in Case study 3.

$$O_{event} = \frac{N_{events}}{N_{non-events}}$$ **Equation 12.**

O_{event} = odds for occurrence of observed phenomenon
N_{events} = number of events of observed phenomenon (part of a whole)
$N_{non-events}$ = number of non-events of observed phenomenon (part of a whole)

Odds are very powerful analytical tool in epidemiology (8,22). Technically we distinguish between odds in different kind of situations (in different types of study design):

- when we are observing the presence of exposure in case-control studies we calculate "exposure-odds",
- when we are observing in a cross-sectional study the frequency of all cases of a disease versus all non-cases we are talking about odds for having a disease in a specified point of time or "prevalence-odds",
- when we are observing in a longitudinal study the occurrence of new cases of disease versus non-occurrence we are talking about odds for getting a disease in a specified period of time or "disease-odds" which are the estimate of risk-odds in the sense of incidence-odds (the concepts of incidence and prevalence are out of the scope of this module, and are discussed in a special module).

At the end we need to stress that the mathematical properties of odds make them advantageous for various uses. Whereas probabilities are restricted to the 0-1 interval, odds can be any nonnegative number. The logarithm of

the odds can therefore be any real number. The natural logarithm of the odds (called the "logit") is relatively widely used in biostatistics and epidemiology (8,18,22).

Application of frequency measures in epidemiology

Both, absolute and relative frequency measures represent the basic tool in epidemiology. They could be classified in three big groups of epidemiologic measures (1,4-7,16,18,19,23-29) (Figure 7):

1. **Frequency measures**

These measures (Figure 7) are also called measures of disease occurrence, measures of occurrence of disease and other health related events, or measures of extent. They are trying to answer to the question how often do happen the observed phenomena (diseases, death etc.) in the population.

2. **Measures of association**

These measures (Figure 7) are also called measures of effect. They are trying to answer to the question why do happen the observed phenomena more often is some population groups than in others. To be able to answer to this question we relate different phenomena to each other. Thus these measures express the extent of association between two (or more) phenomena one of them usually being negative health phenomenon and the other risk factor (putative cause) for the first one.

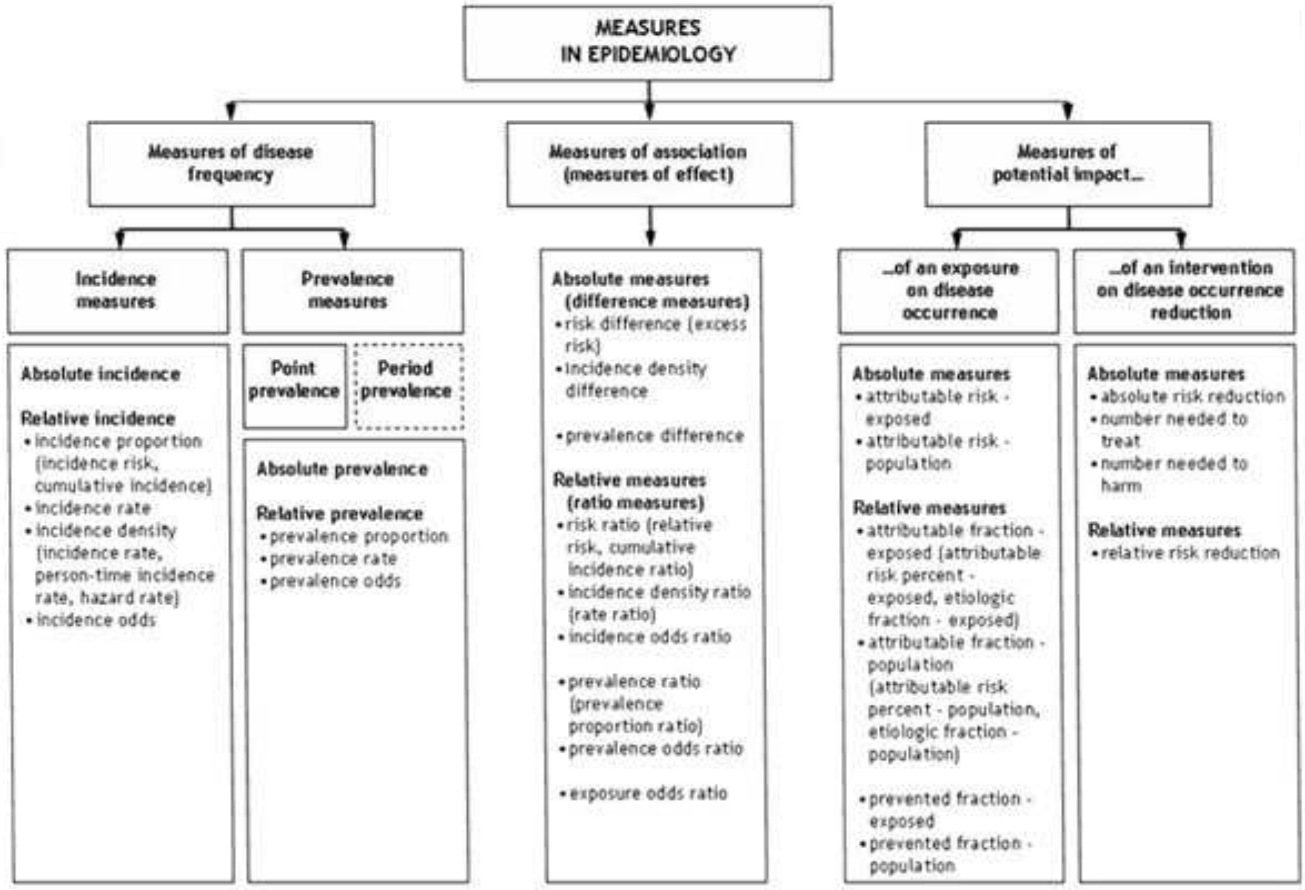
3. **Measures of potential impact**

In this group there are two groups of measures (Figure 7). Measures of the first group express potential impact of risk factor on occurrence of observed health phenomenon among exposed persons or in population. These measures are common in public health. Measures of the second group express potential impact of an intervention on disease occurrence reduction. They are much more common in clinical epidemiology than in public health.

Some of these measures are discussed in details in three separate modules of this book.

In this place we need to give a warning about terminology in the field of epidemiologic measures. A common problem in epidemiology is existence of multiple terms for the same concept. Also, there are instances where a single term is applied to different concepts. The confusion is aggravated by the multitude of terms that have been introduced, with usages that differ from one author to another (18).

Figure 7. An overview of frequently used measures in epidemiology



Case studies

Case study 1: Organizing data

Introduction to data set

For this case study, real data are used. In Slovenia for already several years for the purpose of teaching epidemiologic methods in public health, comprising also statistical methods, data collection which enables learning such methods in much comprehensive way has been created. These are the data collected in the frame of the Perinatal Informational System of Slovenia (PISS) (29), which is considered to be one of the permanent of health data-bases of the highest quality with many years' tradition in the country. It was started in 1987, when collection of perinatal data started according to a common protocol in all fourteen Slovene maternity hospitals.

The basic data material for all epidemiologic and statistical activities is composed of 6,356 statistical units, representing the model of a population. For teaching different epidemiologic and statistical methods, samples of various sizes are then randomly selected from the population database. The data set used in this example is composed of 100 observational units.

Data material for teaching is only a small piece out of the whole collection PISS, prepared especially for this purpose. Safeguard of personal data is assured so that all personal data have been removed, and moreover, the data are selected from the whole collection which shall be used only for the teaching purpose.

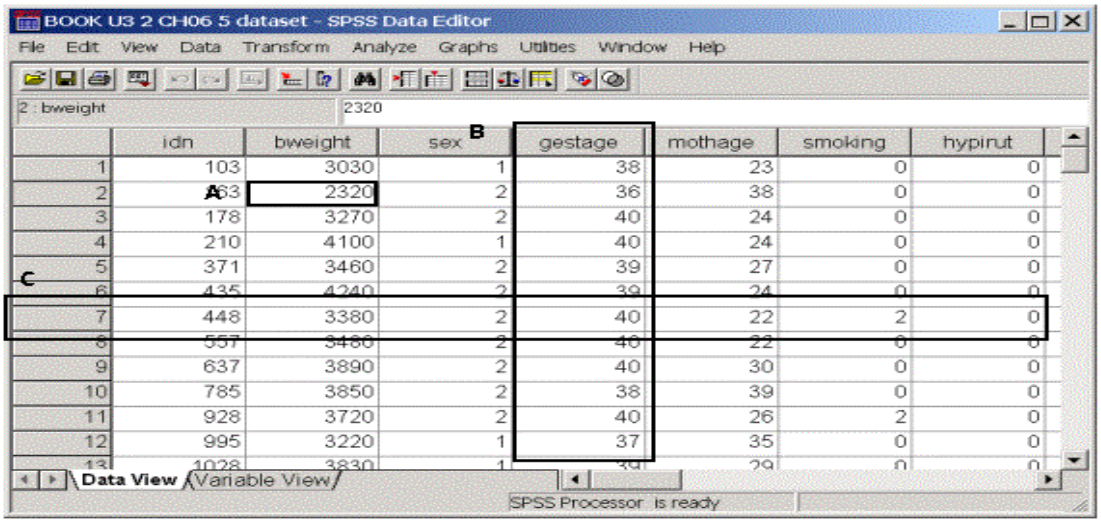
Variables in the data set

In a maternity hospital, data on 100 deliveries were collected. Which characteristics of mother and her child (unit of observation) were observed it is shown in Table 2. Data were organized for description and analysis in a data matrix (Figure 8).

Table 2. Description of variables, their values and codes in demonstrational data set

Column in a data set	Short name of a variable	Information the variable is giving about	Variable values and their codes in data set
1	IDN	unit identification number	
2	BWEIGHT	birth weight (child)	(weight in grams)
3	SEX	sex (child)	1 = boy 2 = girl
4	GESTAGE	gestational age (child)	(age in weeks of pregnancy)
5	MOTHAGE	age at delivery (mother)	(age in years)
6	SMOKING	smoking habits during pregnancy (mother)	0 = no 1 = up to 10 cigarettes/day 2 = 10 cigarettes or more/day
7	HYPIRUT	hyperactivity of uterus (hyper irritable uterus) during pregnancy (mother)	0 = no 1 = yes
8	EBP	elevated blood pressure during pregnancy (mother)	0 = no 1 = yes
9	MEDVIT	consumption of vitamins preparations during pregnancy (mother)	0 = no 1 = yes
10	MEDFE	consumption of iron preparations during pregnancy (mother)	0 = no 1 = yes
11	MEDAB	consumption of antibiotics during pregnancy (mother)	0 = no 1 = yes

Figure 8.Organization of data in data set for Example 1. A – value of variable BWEIGHT for unit 2; B – values for variable GESTAGE for the first twelve units; C - values for the first seven variables (IDN - HYPIRUT) for unit number 7



Case study 2: Statistical description of data

This case study is basing on the same set of data as Case study 1.

Defining variables type

From the Table 2 we can see that variables BWEIGHT and MOTHAGE are numerical variables, while all other variables are categorical. We will now analyze the variable MOTHAGE - age at delivery (mother).

From the Figure 8 could be seen the record for first 12 units and first 7 variables. A record for all seven variables for unit No.7 (ROW 7), and record for variable GESTAGE for first twelve units (COLUMN GESTAGE) is accentuated. The crossing of ROW7 and COLUMN GESTAGE has value 40 (value of variable GESTAGE for unit 7).

Analysis of the variable "MOTHAGE"

In its origin; "AGE" is a numerical continuous variable. Theoretically, the smallest interval between two values of this variable depends on the precision of the device for measuring it, but in practice we are never interested in such precise information so we always reduce it:

- the information is usually limited (reduced) to intervals of 1-year; in observation of different health states the intervals of 1-day, 1-week, 1-month (neonatology, pediatrics) or 5-years (public health) are also used,
- in public health, the information is often reduced even more when we classify (group) the values according to age groups (periods of life) (e.g. babies, preschool children, primary school children, adolescents, adults, aged people); the difference between this and the previous level of reduction of information is that intervals of present level are no more equal while on the previous are,
- the information about age is the most reduced when we divide the whole scale in two parts (e.g. adults of 25 years or more: yes or no).

Variable MOTHAGE is in our case interval numerical variable, with 1-year interval width.

Frequency distribution and histogram

Again, we will use variable MOTHAGE - age at delivery (mother). After arranging of values in ordered series, the next step is to summarize this ordered series in a frequency distribution table. The frequency distribution table of this variable is presented in Figure 9.

In Figure 10, the graphic presentation of frequency distribution for variable MOTHAGE, the histogram, is presented. Next step is to calculate or to determine the typical values of this distribution.

Typical values

The description of the distribution of values of variable MOTHAGE - age at delivery (mother) of example dataset is as follows

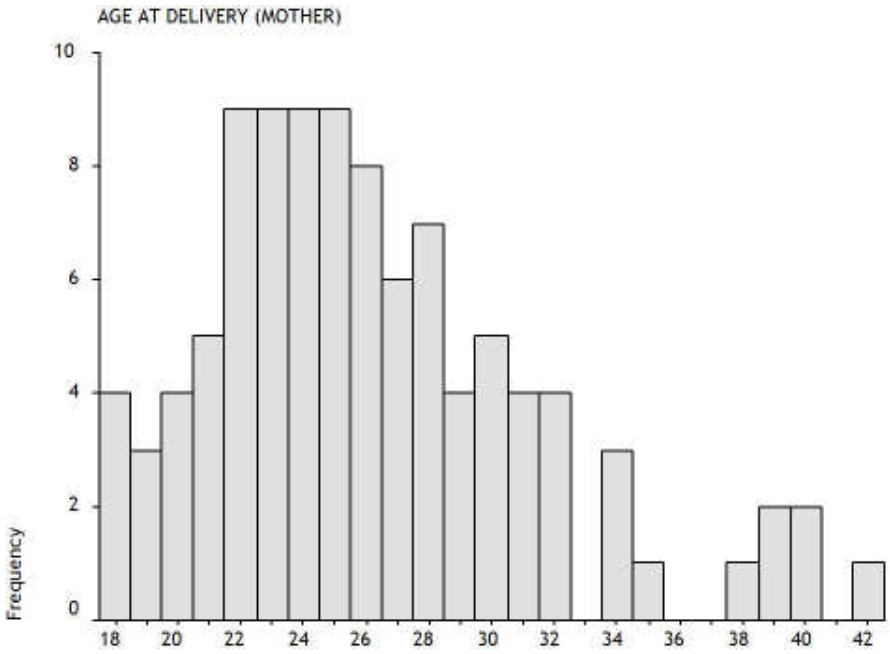
- the distribution is bell shaped (Figures 9 and 10),
- it is slightly asymmetrical (Figure 9 and 10),

Figure 9. An example of a frequency distribution table of variable MOTHAGE - age at delivery (mother) of example dataset (the SPSS statistical programme printout) (30)

AGE AT DELIVERY (MOTHER)

Valid	Frequency	Percent	Valid Percent	Cumulative Percent
18	4	4,0	4,0	4,0
19	3	3,0	3,0	7,0
20	4	4,0	4,0	11,0
21	5	5,0	5,0	16,0
22	9	9,0	9,0	25,0
23	9	9,0	9,0	34,0
24	9	9,0	9,0	43,0
25	9	9,0	9,0	52,0
26	8	8,0	8,0	60,0
27	6	6,0	6,0	66,0
28	7	7,0	7,0	73,0
29	4	4,0	4,0	77,0
30	5	5,0	5,0	82,0
31	4	4,0	4,0	86,0
32	4	4,0	4,0	90,0
34	3	3,0	3,0	93,0
35	1	1,0	1,0	94,0
38	1	1,0	1,0	95,0
39	2	2,0	2,0	97,0
40	2	2,0	2,0	99,0
42	1	1,0	1,0	100,0
Total	100	100,0	100,0	

Figure 10. An example of a histogram of variable MOTHAGE - age at delivery (mother) of example dataset (the SPSS statistical programme printout) (30)



As the number of units is rather large, it could be summarized by mean and standard deviation or median and quartiles/percentiles. The mean and the median value are, because slight asymmetry of the distribution, similar but not the same. The decision which set of typical values to choose is up to investigator. In Figure 11, typical values for the distribution shown in Figure 10 are presented.

Figure 11. An example of set of typical values for the distribution of variable MOTHAGE - age at delivery (mother) of example dataset (the SPSS statistical programme printout) (30)

AGE AT DELIVERY (MOTHER)

N	Valid	100
	Missing	0
Mean		26,19
Median		25,00
Std. Deviation		5,256
Minimum		18
Maximum		42
Percentiles	25	22,25
	50	25,00
	75	29,00

Case study 3: Epidemiologic description of data

Introduction to data set

This case study as well, is based on real data. In Slovenia in 2001 the survey aiming at assessing the prevalence of health behaviours (smoking, nutrition, alcohol consumption, physical activity and traffic safety) was performed for the first time. This survey is conceptually a part of a wider international project in the frame of the Countrywide Integrated Non-communicable Diseases Intervention (CINDI) program, supported by the World Health Organization, CINDI Health Monitor.

The stratified random sample was drawn from the Central Population Registry of the Republic of Slovenia. The sample size was 15,379 with the age range 25-64 years. The sampling was performed by the Statistical Office of the Republic of Slovenia.

The data were collected in late spring 2001 by using a self-administered postal questionnaire, based on the CHM Core Questionnaire (31).

Out of 15,379 inhabitants included in the sample 15,153 were contacted (226 were excluded because of changed domicile, severe illness or death). The response rate was 63.8% (9,666 responses). The respondents did not differ statistically from non-respondents in age distribution or distribution of size of settlements of permanent residence, but the response to the survey was slightly lower among men (47.0%) than among women (53.0%) at a ratio 1:1.1 (according to population data in 2001 the ratio was 1:1). The questionnaires of 9,034 respondents were eligible for analysis (eligibility criteria: sex and age provided by Statistical Office of the Republic of Slovenia).

For the purpose of this module, we have chosen observation of smoking behaviour.

Variables in the data set

In CINDI Health Monitor survey in Slovenia in 2001 (CHMS 2001) 8,904 respondents reported their current smoking status. The answers grouped regarding the sex of respondents are shown in Table 3.

Table 3. Smoking status in CINDI Health Monitor survey in Slovenia in 2001 in total sample and by sex

	SEX		Total
SMOKER	Male	Female	
No	2,931	3,859	6,790
Yes	1,143	971	2,114
Total	4,074	4,830	8,904

Ratios

Proportion and percentage

In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the time of the survey (Table 3). The proportion of smokers could be calculated as a vulgar fraction (Equation 13):

$$\text{Proportion} = \frac{2{,}114}{8{,}904} \qquad \textbf{Equation 13.}$$

or as a decimal fraction (Equation 14):

$$\text{Proportion} = \frac{2{,}114}{8{,}904} = 0.237 \qquad \textbf{Equation 14.}$$

or as percentage (Equation 15):

$$\text{Proportion} = \frac{2{,}114}{8{,}904} \times 100 = 23.7\% \qquad \textbf{Equation 15.}$$

Rate in classic epidemiologic sense

In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the moment of the survey (Table 3). As the survey is a representative of cross-sectional studies the time component is a point in time (a moment) (Equation 16):

$$\text{epidemiologic Rate}_{(\text{at the moment of a survey})} = \frac{2{,}114}{8{,}904} \times 1{,}000 = 237 \qquad \textbf{Equation 16.}$$

The epidemiologic rate (prevalence rate) has value 237 per 1,000 population.

Ratio in a narrow sense

We could express several ratios in a narrow sense using this example:

1. In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the time of the survey and the other 6,790 that they were not (Table 3). The ratio between non-smokers and smokers is (Equation 17):

$$\text{Ratio} = \frac{6{,}790}{2{,}114} = 3.21 \qquad \textbf{Equation 17.}$$

The ratio is 3.21 to 1, what means that in Slovenia in 2001 there were 3.21-times more non-smokers than smokers (or there were 3.21 non-smokers to one smoker).

2. If we now turn the ratio and observe the ratio of smokers to non-smokers we get (Equation 18):

$$\text{Ratio} = \frac{2{,}114}{6{,}790} = 0.31 \qquad \textbf{Equation 18.}$$

The rate is 0.31 to 1, what means that in Slovenia in 2001 there was 0.31 of a smoker to one non-smoker.

3. In CHMS 2001, 2,114 respondents reported that they were smokers at the time of the survey. Among them there were 1,143 men in 971 women (Table 3). The ratio between men and women among smokers was (Equation 19):

$$\text{Ratio} = \frac{1{,}143}{971} = 1.18 \qquad \textbf{Equation 19.}$$

The ratio was 1.18 to 1, what means that in Slovenia in 2001 there were 1.18-times more male smokers than female smokers (or to one female smoker there were 1.18 male smokers).

Probability and odds

Probability

In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the time of the survey (Table 3). The probability for being a smoker at the time of survey could be calculated as:

- a vulgar fraction (Equation 20):

$$p = \frac{2{,}114}{8{,}904} \qquad \textbf{Equation 20.}$$

- as a decimal fraction (Equation 21):

$$p = \frac{2{,}114}{8{,}904} = 0.237 \qquad \textbf{Equation 21.}$$

- or as a percentage (Equation 22):

$$p = \frac{2{,}114}{8{,}904} \times 100 = 23.7\% \qquad \textbf{Equation 22.}$$

The probability for being a smoker at the moment of the survey CHMS 2001 was, expressed as percentage, 23.7%.

Odds

The odds for being a smoker for data presented in Table 3 could be calculated in two different ways.

1. Through calculating first the probability for being a smoker at the moment of the CHMS 2001 survey (also interpreted as risk in an epidemiologic sense - a probability for unfavourable health behaviour in this case) and the probability for being a non-smoker. The probability for being a smoker was (Equation 27):

$$p = \frac{2{,}114}{8{,}904} = 0.237 \qquad \textbf{Equation 27.}$$

while the probability for being a non-smoker was (Equation 28):

$$1 - p = 1 - 0.237 = 0.763 \qquad \textbf{Equation 28.}$$

The odds for being a smoker can be calculated now as (Equation 29):

$$O = \frac{0.237}{1 - 0.237} = \frac{0.237}{0.763} 0 = 0.31 \qquad \textbf{Equation 29.}$$

2. Through algebraic cancellation of total number of possible events from calculation of odds:

$$O = \frac{\frac{2{,}114}{8{,}904}}{1 - \frac{2{,}114}{8{,}904}} = \frac{\frac{2{,}114}{8{,}904}}{\frac{8{,}904}{8{,}904} - \frac{2{,}114}{8{,}904}} = \frac{\frac{2{,}114}{8{,}904}}{\frac{6{,}790}{8{,}904}} = \frac{2{,}114}{6{,}790} = 0.31 \qquad \textbf{Equation 30.}$$

The odds for being a smoker versus non-smoker are 0.31. This means that in Slovenia in 2001 there was 0.31 of a smoker to one non-smoker, or the ratio non-smokers to smoker is 1 to 0.31. This is exactly the same result as in Equation 29.

Exercises

Task 1: Statistical description of data

1. From the table with description of the basic data set (Table 2) find out how many variables are in the data set, their names and which information they contain.
2. Find out which of the variables could play the role of "the effect" and which ones the role of "the cause".
3. Carefully read the following statements and determine how many variables are included and which could be their values:
   - after the fractures children recover faster than adults,
   - men with inflammation of joints differ in response to therapy from women,
   - men more often get lung complications after the heart operation then women.
4. In the table with raw data set (APPENDIX, Table A1) find out how many units are there in this sample.
5. Find out what is the unit under study.
6. Enter the data for the first 20 units in data matrix.
7. For the variable MOTHAGE make the ordered series.
8. Make the frequency distribution table (with absolute and relative frequencies in percentages) for this variable, too.
9. Find out for the same variable if any value exists with absolute frequency equal to 0.
10. Make the frequency distribution table for the same variable (MOTHAGE) also with your statistical program and compare it with the table you made in Exercise 8.
11. Make frequency distribution tables with your statistical program also for the following variables: BWEIGHT, SEX, GESTAGE, SMOKING and HYPIRUT.
12. For variables BWEIGHT, SEX, GESTAGE, MOTHAGE, SMOKING and HYPIRUT find out:
    - ordinality of their values
    - how many values can you find for each one in a frequency distribution table,
    - find out if could you classify these values as continuous or discrete,
    - classify each variable according to type,
    - for numerical variables (continuous and discrete) find out from the frequency distribution table the lowest, the highest and the most frequent value/s (if there are more than one, list all of them),
13. Find out from the frequency distribution of variables MOTHAGE and GESTAGE for each one:
    - where the density of distribution is the highest,
    - is the distribution symmetrical or not and if it is not, to which direction it is skewed,
    - where would you locate the centre of the distribution.

    Compare both frequency distributions.

14. For variables MOTHAGE and GESTAGE draw a histogram using the absolute frequency from the frequency distribution table; consider also all values with the frequency equal to 0.
15. Draw a histogram for variable BWEIGHT, too; as there are many different values, group observations in intervals of 250gr of width in order to get a better visual impression of the observed distribution.
16. Find out in histogram for each variable:
    - is the distribution bell shaped or not,
    - is the distribution symmetrical or not and if it is not, to which direction it is skewed,
    - where would you locate the centre of the distribution.

    Compare all three histograms.
17. Compare these conclusions with your conclusions in Exercise 12.
18. Make the histograms also with your statistical program and compare them to histograms you made by yourself.
19. For attributable variables SEX, SMOKING and HYPIRUT draw the ordinary bar charts by yourself.
20. Make the bar charts for these variables also with your statistical program and compare them to bar charts you made by yourself in Exercise 18.
21. For numerical variables BWEIGHT, GESTAGE and MOTHAGE make a decision which typical values would be the most appropriate for summarizing the features of distribution of their values and prove it.
22. For these variables determine the typical values by the means of your statistical programme.
23. What can you say about the distribution of a variable for which the mean and the median differ significantly
24. Find out if it is sensible to determine typical values for any type of attributable variables
25. Carefully read the following statement and decide if it is true or false:
    In attributable variable with only two values, 0 and 1, the proportion of units with value 1 equals to hidden arithmetic mean of this type of variable
26. Check your decision with statistical program: calculate the proportion of units with value 1 and the arithmetic mean of the variable HYPIRU.

Task 2: Epidemiologic description of data

This task is based again on PISS data (28).

Table 3. Description of newly designed variables, created from original ones, their values and codes in demonstrational data set

Short name of a variable	Original variable (cut-off point)	Information the variable is giving about	Variable values and their codes in data set
LBW	BWEIGHT (2500 g)	Low birth weight (2500 g or less)	0 = no 1 = yes
SMOKING1	SMOKING	Smoking during pregnancy	0 = no 1 = yes

Frequency of low birth weight (LBW) was observed in two groups of newborns according to smoking of mother during pregnancy. The results are presented in Table 4. All following exercises are referring to this table.

Table 4. Frequency of low birth weight (LBW) in newborns in two groups divided according to smoking of mother during pregnancy based on PISS data (29)

Low birth weight	Smoking of mother during pregnancy		Total
	No	Yes	
No	**558**	**191**	749
Yes	35	16	51
Total	593	207	800

1. Calculate following ratios in which numerator is included in denominator:
    - a proportion of LBW newborns as a decimal fraction in the total sample of newborns,
    - a percentage of LBW newborns in the total sample of newborns,
2. Calculate the ratio in a narrow sense of LBW in the total sample of newborns.
3. Calculate probability of LBW in the total sample of newborns, in the group of smoking mothers, and in the group of non-smoking mothers during pregnancy.
4. Calculate odds of LBW in the total sample of newborns, in the group of smoking mothers, and in the group of non-smoking mothers during pregnancy.

References

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.

3. Timmreck TC. An introduction to epidemiology. 3rd ed. Sudbury, MA: Jones&Bartlett Publishers; 2002.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Co.; 2001.
5. Centers for Disease Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics. Atlanta: Centers for Disease Control and Prevention. Available from URL: www.cdc.gov/osels/scientific_edu/ss1978/SS1978.pdf. Accessed: June 6, 2013.
6. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maidenhead: Open University Press, McGraw-Hill Education; 2005.
7. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999.
8. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1993.
9. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
10. ISI - International Statistical Institute. ISI Multilingual Glossary of Statistical Terms [homepage on the Internet]. Available from URL: http://isi.cbs.nl/glossary.htm. Accessed: June 6, 2013.
11. StatSoft, Inc. Electronic statistics textbook. Tulsa, OK: StatSoft; 2007. Available from URL: http://www.statsoft.com/textbook/stathome.html. Accessed: June 6, 2013.
12. Kosmelj B, Arh F, Dobersek-Urbanc A, Ferligoj A, Omladic M. Statistical dictionary (in Slovene). Ljubljana: Statisticno drustvo Slovenije: Studentska zalozba; 2002.
13. Swinscow TDV. Statistics at square one. Ninth Edition. London: BMJ Publishing Group; 1997. Available from URL: http://www.bmj.com/collections/statsbk/. Accessed: June 6, 2013.
14. Hicks C. Research for physiotherapists: project design and analysis. Edinburgh: Livingstone; 1995.
15. Friis RH, Sellers TA. Epidemiology for public health practice. Fourth edition. Sudbury, MA: Jones&Bartlett Publishers; 2009.
16. Kue Young T. Population health: concepts and methods. Second edition. Oxford: Oxford University Press; 2005.
17. Babus V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
18. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 6, 2013.
19. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.
20. Easton VJ, McColl JH. Statistics glossary. V 1.1. Probability [homepage on the Internet]. Available from URL: http://www.stats.gla.ac.uk/steps/glossary/alphabet.html. Accessed: June 6, 2013.
21. Organization for Economic Co-operation and Development - OECD. Glossary of statistical terms. Probability [homepage on the Internet]. Available from URL: http://stats.oecd.org/glossary/detail.asp?ID=3794. Accessed: June 6, 2013.
22. Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 2000.
23. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
24. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
25. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
26. Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005.
27. Pearce N. Effect measures in prevalence studies. Environ Health Perspect. 2004;112:1047-50.
28. Jenicek M. Foundations of evidence-based medicine. London: Parthenon publishing; 2003.
29. Institute of Public Health of the Republic of Slovenia. Perinatal Informational system of the Republic of Slovenia. Available from URL: http://www.ivz.si/Mp.aspx?ni=46&pi=5&_5_id=56&_5_PageIndex=1&_5_groupId=185&_5_newsCategory=&_5_action=ShowNewsFull&pl=46-5.0. June 6, 2013.
30. SPSS Base 7.5 for Windows. User's Guide. SPSS Inc. Chicago; 1997.
31. Zaletel-Kragelj L, Fras Z, Maucec Zakotnik J, editors. Risky behaviours related to health and selected health conditions in adult population of Slovenia: results of Slovenia CINDI Health Monitor Survey 2001. II. Risky behaviours (in Slovene). Ljubljana: CINDI Slovenija; 2004. p.107-564.

Recommended readings

1. Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1993.
2. Babus V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
3. Centers for Disease Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics. Atlanta: Centers for Disease Control and Prevention. Available from URL: www.cdc.gov/osels/scientific_edu/ss1978/SS1978.pdf. Accessed: June 5, 2013.
4. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999.
5. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
6. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
7. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.

Appendix

Table A1. Raw perinatal data. Perinatal Informational System of Slovenia (PISS) (29)

	IDN	BWEIGHT	SEX	GESTAGE	MOTHAGE	SMOKING	HYPIRUT	HBP	MEDVIT	MEDFE	MEDAB
1.	103	3030	1	38	23	0	0	0	0	1	0
2.	163	2320	2	36	38	0	0	0	0	1	0
3.	178	3270	2	40	24	0	0	0	0	0	0
4.	210	4100	1	40	24	0	0	0	0	0	1
5.	371	3460	2	39	27	0	0	0	1	0	0
6.	435	4240	2	39	24	0	0	0	0	0	0
7.	448	3380	2	40	22	2	0	0	0	0	0
8.	557	3480	2	40	22	0	0	0	1	1	0
9.	637	3890	2	40	30	0	0	0	0	0	0
10.	785	3850	2	38	39	0	0	0	0	0	1
11.	928	3720	2	40	26	2	0	0	0	1	0
12.	995	3220	1	37	35	0	0	0	1	1	1
13.	1028	3830	1	39	29	0	0	0	1	1	0
14.	1034	3680	2	40	21	0	0	0	0	0	0
15.	1048	3160	2	37	23	0	0	0	0	0	0
16.	1142	3250	1	39	29	0	0	0	0	0	0
17.	1143	4130	1	40	26	0	0	0	0	1	1
18.	1171	2980	1	39	24	0	0	0	1	0	0
19.	1209	4900	1	41	26	1	0	0	0	0	0
20.	1258	1880	2	37	30	0	0	0	0	0	0
21.	1365	2100	2	37	26	1	0	0	0	0	0
22.	1397	5000	2	42	28	0	0	0	1	1	0
23.	1424	3430	2	39	25	0	1	0	0	1	0
24.	1426	3590	1	38	42	0	0	0	0	0	0
25.	1472	3680	2	39	23	0	0	0	0	0	0
26.	1473	3320	2	41	23	0	0	0	0	0	0
27.	1576	3560	1	39	19	0	0	0	0	0	0
28.	1588	3430	1	41	31	0	0	0	0	1	0
29.	1604	1840	2	36	27	0	0	1	0	0	1
30.	1620	3170	1	40	19	0	0	0	0	1	0
31.	1642	3740	1	41	32	0	0	0	0	0	1
32.	1706	3130	2	41	28	0	0	0	1	1	0
33.	1808	3460	2	39	22	0	0	0	1	1	1
34.	2021	3710	1	40	28	1	0	0	0	1	0

Table A1. Cont.

	IDN	BWEIGHT	SEX	GESTAGE	MOTHAGE	SMOKING	HYPIRUT	HBP	MEDVIT	MEDFE	MEDAB
35.	2031	3120	1	41	27	0	0	0	0	0	0
36.	2096	2200	2	34	22	2	0	0	0	0	0
37.	2166	1530	1	32	25	0	0	0	1	1	1
38.	2264	2820	1	40	23	1	0	0	0	1	0
39.	2269	3880	2	38	27	1	0	0	0	0	0
40.	2346	3870	1	39	24	0	0	0	0	0	1
41.	2499	3450	2	40	29	1	0	0	0	0	0
42.	2632	2770	1	35	18	0	0	0	0	1	0
43.	2668	3480	1	41	25	0	0	0	1	1	0
44.	2747	3300	1	38	19	0	0	0	0	0	0
45.	2786	3920	1	39	25	0	0	0	0	1	0
46.	2799	4140	2	40	28	0	0	0	0	1	0
47.	2871	3470	2	41	23	1	0	0	1	1	0
48.	2965	3700	2	40	31	0	0	0	0	1	0
49.	3092	2420	2	37	29	0	0	0	1	0	0
50.	3127	3400	1	38	30	0	0	0	0	0	0
51.	3156	2770	2	38	23	1	0	0	0	0	0
52.	3170	3590	2	40	21	0	0	0	1	1	0
53.	3220	3800	1	40	22	0	1	0	1	1	1
54.	3286	3370	1	40	32	2	0	0	0	1	1
55.	3314	3200	2	39	26	1	0	0	0	0	0
56.	3333	3480	2	38	26	0	0	0	0	1	0
57.	3379	2920	1	38	34	1	1	1	0	0	0
58.	3417	3850	1	40	25	0	0	0	0	1	0
59.	3430	3160	2	41	20	2	0	0	0	0	0
60.	3469	3500	1	39	30	0	0	0	1	0	0
61.	3471	2970	2	39	27	0	0	0	1	0	0
62.	3498	3640	2	39	34	0	0	0	0	0	0
63.	3501	2440	2	36	18	0	0	0	0	0	0
64.	3567	2660	1	39	28	0	0	0	0	0	0
65.	3604	3210	2	40	24	1	0	0	0	0	0
66.	3621	3260	1	39	20	0	0	0	1	1	0
67.	3732	3730	1	41	23	0	0	0	1	0	0
68.	3851	3040	2	39	22	0	0	0	1	1	0
69.	3918	2940	1	38	25	0	0	0	0	1	0
70.	3923	2920	2	40	31	0	0	0	1	0	1

Table A1. Cont.

	IDN	BWEIGHT	SEX	GESTAGE	MOTHAGE	SMOKING	HYPIRUT	HBP	MEDVIT	MEDFE	MEDAB
71.	4019	3460	2	39	40	0	0	0	1	0	0
72.	4034	3360	2	40	24	0	0	0	0	1	1
73.	4145	3140	1	41	25	0	1	0	0	0	0
74.	4193	3800	2	37	40	0	1	0	0	1	0
75.	4206	3350	1	39	25	1	0	0	1	0	0
76.	4209	3350	2	40	24	1	0	0	1	1	0
77.	4386	3810	2	41	32	0	0	0	0	1	0
78.	4421	2420	1	35	30	0	0	0	0	0	0
79.	4522	3050	2	39	22	1	0	0	0	0	0
80.	4598	2840	1	39	39	0	1	0	1	0	0
81.	4672	4170	1	41	25	0	0	0	1	0	0
82.	4944	2410	2	38	26	0	0	0	0	1	0
83.	4957	3780	1	40	22	0	0	0	0	0	0
84.	5002	4130	1	42	21	1	0	0	0	0	0
85.	5122	2830	1	39	18	0	0	0	0	0	0
86.	5249	3960	2	40	18	0	0	0	0	0	0
87.	5409	3280	2	39	21	0	0	0	0	1	0
88.	5433	3020	2	37	28	0	1	0	0	1	0
89.	5445	1270	1	30	22	0	0	0	0	0	0
90.	5564	3100	1	40	21	0	1	0	0	1	1
91.	5656	3870	1	38	23	0	0	0	0	1	0
92.	5678	3470	1	40	24	2	0	0	0	0	0
93.	5761	4000	1	40	26	0	0	0	0	0	0
94.	5820	2960	2	40	20	0	0	0	1	0	0
95.	5862	3670	1	40	27	0	0	0	0	1	0
96.	5871	3400	2	38	32	0	0	0	1	1	0
97.	5887	2840	1	38	20	0	0	0	1	0	0
98.	5907	3060	1	38	34	0	0	0	1	1	0
99.	6013	3230	2	39	31	0	1	0	0	1	0
100.	6131	3230	2	40	28	0	0	0	0	0	1

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Frequency measures: Prevalence and incidence**
Module: 2.3	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Lijana Zaletel-Kragelj Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Frequency measures, incidence, prevalence.
Learning objectives	After completing this module students should: know and understand the difference between prevalence and incidence measures of frequency, be familiar with different types of prevalence measures, and know how to calculate them; be familiar with different types of incidence measures, and know how to calculate them.
Abstract	Frequency measures are quantities that express frequency of health phenomena. Prevalence and incidence are two main groups of frequency measures. The most distinctive difference between these two groups is that by prevalence measures we are observing the transversal section through the situation of phenomenon under observation at designated time (e.g. in a point of time) while by incidence measures we are observing its dynamics (by performing regular observation of breaking out of new cases of phenomenon under observation during every of specified equal time periods) in a specified population.
Teaching methods	An introductory lecture gives the students first insight in features and types of frequency measures. The theoretical knowledge is illustrated by case studies. After introductory lectures students first carefully read the theoretical background of this module and complement their knowledge with recommended readings. Afterwards they on provided data set perform extensive tasks on calculation of different types of measures. They are stimulated to compare results with results of each other and discuss the differences.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on calculated measures in which detailed description of process of calculation is described.

FREQUENCY MEASURES: PREVALENCE AND INCIDENCE

Lijana Zaletel-Kragelj

Theoretical background

Introduction to frequency measures

Prevalence and incidence are two main groups of frequency measures in epidemiology. We should be conscious that there are in fact two families of measures under each term although we are frequently talking about only two measures. This is the reason of great deal of misunderstanding and misinterpretation of frequency measures. The most distinctive difference between these two groups of measures is the fact that by prevalence measures we are observing the transversal section through the situation of phenomenon under observation most usually at a designated time (e.g. in a point of time), while by incidence measures we are observing its dynamics (by performing regular observation of breaking out of new cases of phenomenon under observation during every of specified equal time periods) in a specified population (1-22). General equations for these two families of measures are as follows (Equation 1 and Equation 2) (1-3):

$$P = \frac{N_{d+\,all\,cases(dt)}}{N_{all\,persons(dt)}} \qquad \textbf{Equation 1.}$$

P = prevalence
$N_{d+\,all\,cases\,(dt)}$ = number of all persons with the disease under observation (cases) at designated time
$N_{all\,persons\,(dt)}$ = number of all persons under observation at designated time

$$I = \frac{N_{d+\,new\,cases(gp)}}{N_{all\,persons\,at\,risk(bgp)}} \qquad \textbf{Equation 2.}$$

I = incidence
$N_{d+\,new\,cases\,(gp)}$ = number of new cases of the disease under observation during a given period
$N_{all\,persons\,at\,risk\,(bgp)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given period

Both families could be theoretically classified regarding various features what will be discussed later.

The process of explanation of differences between both families of measures and the differences between measures inside both families will be illustrated in the case studies.

Before starting with explanation of concepts of incidence and prevalence, both extremely important epidemiologic concepts, it could be worthy to emphasize that a common problem in epidemiology is the existence of multiple terms for the same concept. Also, there are instances where a single term is applied to different concepts. The confusion is aggravated by the multitude of terms that have been introduced, with usages that differ from one author to another (21)

At the beginning it could be useful for students to emphasize that there exist different frequency measures on one hand and different study designs on the other. Although we could on a theoretical level show that we have several analogous frequency measures, not all of them are of "common sense" and are not meaningful in all situations. One should be aware that it is important to use the measure that is most appropriate for the current task (22). Since, we think that one can choose the most appropriate measure if he/she fully understands the differences between them.

Most of the textbooks on epidemiologic methods first concentrate on incidence measures, and only afterwards on prevalence measures. We will do the opposite way since we think that the concept of prevalence might be easier to understand.

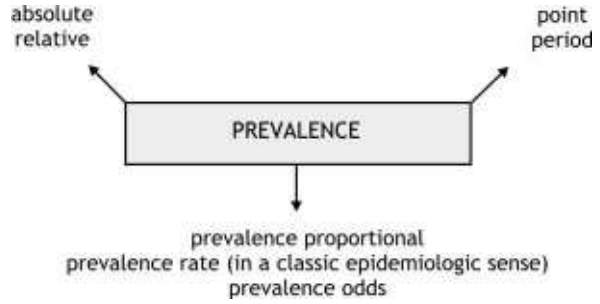
Prevalence

Definition

Prevalence is a common term for a group of measures which are quantifying the state (situation) of a given health phenomenon (e.g. a disease, a disorder, an unhealthy health behaviour, etc.) at a designated time (at a specified moment, or at any time during a specified period) irrespective of wheather the cases of observed phenomenon are old or new (1-6,12-17,21,22). Prevalence indices are measuring, thus, the burden of disease or any other health condition in a population or its power (praevaleo lat. to be very strong) at a designated time. A special problem could be the fact that the term "prevalence" could denote various prevalence measures, mostly as a synonym for "prevalence rate" (in a classic epidemiologic sense).

Prevalence measures could be classified according to various characteristics. Three different classifications are summarized in Figure 1.

Figure 1. Classifications of prevalence measures according to various characteristics



Classification of prevalence measures according to type of frequency measure upon which are based

According to the type of frequency measure upon which are based, we distinguish between absolute and relative prevalence.

1. Absolute prevalence, prevalent number or prevalence

This measure is simply a number of all cases of observed phenomenon at a designated time (at a specified moment, or at any time during a specified period).

Absolute prevalence frequency measures are important in health care planning (e.g. number of hospital beds needed for treatment of certain group of health states). Calculation of this measure in practice is presented in "Case study 1".

2. Relative prevalence

Relative prevalence are several measures based upon relative frequency (based on different measures of relative frequency) - prevalence as a proportion, prevalence as a rate (in a classic epidemiologic sense), and prevalence as odds. All three measures will be discussed later in detail, only the first one, the prevalence as a proportion, is presented in this first place to be contrasted to absolute prevalence (Equation 3):

$$P_{rel\,(proportion)} = \frac{N_{d+\,all\,cases(dt)}}{N_{all\,persons(dt)}} \qquad \textbf{Equation 3.}$$

$P_{rel\,(proportion)}$ = relative prevalence as a proportion
$N_{d+\,all\,cases\,(dt)}$ = number of all persons with the disease under observation (cases) at designated time
$N_{all\,persons\,(dt)}$ = number of all persons under observation at designated time

Prevalence proportion is a probability of having a disease at a time t, or a probability that an individual will be a case at time t (7).

Relative prevalence frequency measures are important in comparisons (e.g. between two or more population groups, between two or more populations, etc.)

Calculation of this measure in practice is presented in "Case study 1".

Classification of prevalence measures according to the type of a designated time of observation

According to the type of a designated time of observation, which could be a specified moment, or a specified period, we distinguish between point and period prevalence. When used without any qualification, the term prevalence refers usually to point prevalence.

1. Point prevalence (1-3)

It is the measure which expresses the burden of the disease under observation at a specified point in time. It could be absolute (absolute point prevalence), or relative (relative point prevalence). A specified point in time could be:

- a specified point in calendar time (e.g. calendar day, calendar week, calendar month), or
- a specified point in the natural course of the disease (e.g. the point of onset of the symptoms), or a specified event that may be associated with or produce changes in human health (e.g. a specified event in a life course, which could be different by calendar time for any of individuals under observation, like onset of puberty, menopause, the beginning of retirement, third post-operative day, etc.).

Relative point prevalence could be expressed as a proportion, rate (in a classic epidemiologic sense), or odds. Point prevalence expressed as prevalence proportion is a probability for having a disease under observation at a specified point in time (7,22), and it could be calculated as follows (Equation 4):

$$ {}_{point}P_{rel\,(proportion)} = \frac{N_{d+\,all\,cases\,(point\,in\,time)}}{N_{all\,persons\,(point\,in\,time)}} $$ **Equation 4.**

${}_{point}P_{rel\,(proportion)}$ = point prevalence as proportion (probability) for having a disease at a specified point in time
$N_{d+\,all\,cases\,(point\,in\,time)}$ = number of all persons with the disease under observation (cases) at a specified point in time
$N_{all\,persons\,(point\,in\,time)}$ = number of all persons under observation at a specified point in time

Calculation of this measure in practice is presented in Case study 1.

2. Period prevalence (1-3,7,10)

Another prevalence measure is period prevalence which is much less frequently used as point prevalence. It is the measure which expresses the probability that an individual in a population will be a case any time during a period of time (7,10), and it could be calculated as a ratio (not a proportion) as follows (Equation 5):

$$ {}_{period}P_{rel} = \frac{N_{d+\,all\,cases\,(period\,of\,time)}}{N_{all\,persons\,(period\,of\,time)}} = \frac{N_{d+_0} + N_{d+\,new\,cases\,(period\,of\,time)}}{N_{all\,persons\,(period\,of\,time)}} $$ **Equation 5.**

${}_{period}P_{relative}$ = period prevalence as a probability for having a disease at any time during a specified period
$N_{d+\,all\,cases\,(period\,of\,time)}$ = number of all persons with the disease under observation (cases) at any time during a specified period
$N_{all\,persons\,(period\,of\,time)}$ = number of all persons in the population for this same period
$N_{d+\,0}$ = number of persons with the disease under observation (cases) at the beginning of the specified period
$N_{d+\,new\,cases\,(period\,of\,time)}$ = number of new cases of the disease under observation during a specified period

This measure requires the assumption of a stable dynamic population for estimation (7,10). If the study population is unstable, this measure has little practical value.

Period prevalence is most often used in situations when the exact time of the onset of a phenomenon under observation for individual cases is not known (7). Calculation of this measure in practice when the assumption of a fixed cohort is met is presented in "Case study 1".

In continuation, we will discuss in detail only the point prevalence.

Classification of relative point prevalence measures according to the type of relative frequency measure upon which are based

According to the type of relative frequency measure upon which is based a point prevalence measure, we distinguish between relative point prevalence as a proportion, as a rate, and as odds.

1. Measures in which the numerator is included in the denominator

Prevalence proportion (7,22)

This measure expresses the probability of having a disease at a designated time under observation. We have already presented this measure (Equation 4), but since here we introduce the notation usually used in epidemiologic textbooks, we repeat it as a new equation (Equation 6):

$$ P = \frac{N_{d+\,all\,cases\,(point\,in\,time)}}{N_{all\,persons\,(point\,in\,time)}} $$ **Equation 6.**

P = prevalence proportion
$N_{d+\,all\,cases\,(point\,in\,time)}$ = number of all persons with the disease under observation (cases) at designated time
$N_{all\,persons\,(point\,in\,time)}$ = number of all persons under observation at designated time

Calculation of this measure in practice is presented in "Case study 1".

Prevalence rate (in a classic epidemiologic sense) (1,2,12,23)

This measure is very similar to the first one. The only difference is that it has additional components – the multiplier and a time component. It is a rate in a classic epidemiologic sense and, when a point prevalence, it is calculated as follows (Equation 7):

$$ PR = \frac{N_{d+\,all\,cases\,(point\,in\,time)}}{N_{all\,persons\,(point\,in\,time)}} \times K $$ **Equation 7.**

PR = prevalence rate
$N_{d+\,all\,cases\,(dt)}$ = number of all persons with the disease under observation (cases) at designated point in time
$N_{all\,persons\,(dt)}$ = number of all persons under observation at designated point in time
K = multiplier (100, 1000, 10.000, 100.000…)

Calculation of this measure in practice is presented in "Case study 1".

2. **Measure in which the numerator is not included in the denominator**

Prevalence odds (14,22)

Prevalence odds are the ratio of the probability of having a disease to that of not having it at a point in time, or when through algebraic cancellation of total number of possible events the reduction is performed, the ratio of cases to non-cases of the disease under observation at a point in time. Prevalence odds could be calculated as follows (Equation 8):

$$PO = \frac{N_{d+\,all\,cases\,(dt)}}{N_{d-\,all\,cases\,(dt)}} \qquad \textbf{Equation 8.}$$

PO = prevalence odds

$N_{d+\,all\,cases\,(dt)}$ = number of all persons with the disease under observation (cases) at designated time

$N_{d-\,all\,cases\,(dt)}$ = number of all persons without the disease under observation (non-cases) at designated time

Calculation of this measure in practice is presented in "Case study 1".

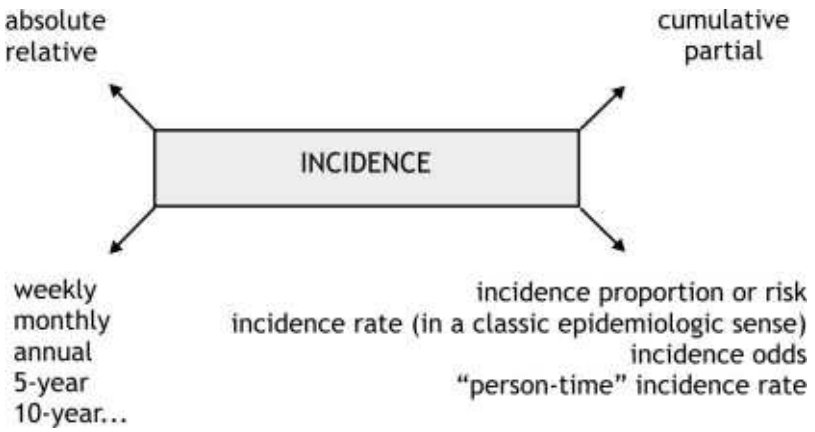
Incidence

Incidence is a common term for a group of measures which are quantifying a break out of new cases of a health phenomenon (e.g. a disease) under observation (incido in morbum lat. to fall ill) during a specified period in a specified group of persons (e.g. population) (1-6,12-17,21,22). A special problem is that the term "incidence" is used to denote various incidence measures.

By performing regular observation of breaking out of new cases of phenomenon under observation during every of specified equal time periods we are observing its dynamics in a specified population.

Incidence measures could be classified according to various characteristics. Four different classifications are summarized in Figure 2.

Figure 2. Classifications of incidence measures according to various characteristics



Classification of incidence measures according to type of frequency measure upon which are based

According to type of frequency measure upon which are based, we distinguish between absolute and relative incidence. The later is more frequently used than the first. When used without any qualification, the term incidence refers usually to absolute incidence, though sometimes is used to mean relative incidence, mostly incidence rate.

1. **Absolute incidence, incident number or incidence (1)**

This measure is simply a number of new cases of observed phenomenon during a specified period in a specified group of persons (e.g. population).

Absolute incidence frequency measures, similarly as absolute prevalence frequency measures, are important in health care planning.

Calculation of this measure in practice is presented in "Case study 2".

2. **Relative incidence**

Relative incidence are several measures based upon relative frequency (based on different measures of relative frequency) - incidence as a proportion, as a rate, as odds, and as an incidence density. All four measures will be discussed later in detail, only the first one, the incidence proportion, is presented in this first place to be contrasted to absolute incidence (Equation 9):

$$I_{rel\,(proportion)} = \frac{N_{d+\,new\,cases\,(gp)}}{N_{all\,persons\,at\,risk\,(bgp)}} \qquad \textbf{Equation 9.}$$

$I_{rel\,(proportion)}$ = relative incidence as proportion (proportion of candidate individuals who developed the disease during the given period)

$N_{d+\,new\,cases\,(gp)}$ = number of new cases of the disease under observation during a given period

$N_{all\,persons\,at\,risk\,(bgp)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given period

Relative prevalence frequency measures are important in comparisons (e.g. between two or more population groups, between two or more populations etc.)

Calculation of this measure in practice is presented in "Case study 2".

Classification of incidence measures according to the fact if the measure expresses the incidence in total observation time or in several parts of it

Although in epidemiology the term "cumulative incidence" is commonly referring to incidence proportion, its intrinsic meaning is referring to cumulation of something (similarly as in statistics). According to this characteristic we distinguish between cumulative and partial, usually annual incidence. The later is the most frequently used measure among possible partial measures.

1. Cumulative incidence

This measure is the number or proportion of individuals under observation in which the onset of observed disease was registered during the entire specified period of observation. Usually it is expressed as a proportion, and it is calculated as follows (equation 10):

$$_{cum}I_{rel\,(proportion)} = \frac{N_{d+\,new\,cases\,(gp)}}{N_{all\,persons\,at\,risk\,(bgp)}} \qquad \textbf{Equation 10.}$$

$_{cum}I_{rel\,(proportion)}$ = cumulative incidence as proportion (proportion of candidate individuals who developed the disease during the entire given period)

$N_{d+\,new\,cases\,(gp)}$ = number of new cases of the disease under observation during a given period

$N_{all\,persons\,at\,risk\,(bgp)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given period

The period of observation (the beginning and the end of the period) has to be exactly stated. The beginning could be based upon calendar time or upon some event in a life-course of individuals under observation (the time of the diagnosis of the disease under observation, or exposure to an agent). This interval is generally the same for all members of the group of individuals under observation, which is true only for fixed cohorts. When withdrawals are present, calculation of this measure is more complicated (7,14,22). Usually, in cohort studies, there are several losses of individuals under observation from follow-up. This is the situation in which the occurrence of the event of interest is uncertain because of different reasons. A situation in which the time-to-event is unknown is called censoring (24). Detailed discussion on this issue is beyond the scope of this module, and is being worked out in a separate module in this book.

Calculation of this measure in practice is presented in "Case study 2".

2. Partial incidence

Total period of observation could be split into several parts in order to get a more correct estimate of incidence, especially the when frequency is varying over time. Annual incidence is usually representative of a partial incidence (one should note that annual incidence could be also a cumulative incidence if a course of a phenomenon under observation is relatively rapid). This measure is the number or proportion of individuals under observation in which the onset of observed disease was registered during the 1-year period. If it is expressed as a proportion, it could be calculated as follows (Equation 11):

$$_{ann}I_{rel\,(proportion)} = \frac{N_{d+\,new\,cases\,(1\text{-}year\,period)}}{N_{all\,persons\,at\,risk\,(beginning\,of\,1\text{-}year\,period)}} \qquad \textbf{Equation 11.}$$

$_{ann}I_{rel\,(proportion)}$ = annual incidence as a proportion (proportion of candidate individuals who developed the disease during the given period)

$N_{d+\,new\,cases\,(1\text{-}year\,period)}$ = number of new cases of the disease under observation during1-year period

$N_{all\,persons\,at\,risk\,(beginning\,of\,1\text{-}year\,period)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given 1-year period

Calculation of this measure in practice is presented in "Case study 2".

Classification of relative incidence measures according to the type of relative frequency measure upon which are based

According to type of relative frequency measure upon which are based relative incidence measures, we distinguish between relative incidence as a proportion, as a rate, as odds, and as an incidence density.

1. Measures in which the numerator is included in the denominator

Incidence proportion (21)

Incidence proportion is a proportion of individuals under observation who developed the disease under observation during a period of observation out of all individuals under observation who were free of disease at the beginning of the specified period of observation (but, at risk for getting the disease). Here we need to introduce two very important terms: "risk" and "cumulative incidence" (1,2,7,9,10,21). Frequently, it seems that risk and cumulative incidence are the same measure, although this is true under very restricted conditions (7). Both terms are closely related to the incidence proportion.

Risk is defined as the probability that a disease-free individual is developing a disease under observation over a specified period, conditional on that the same individual is not dying from any other disease during the period (7). Thus, risk is a conditional probability, with values varying between zero and one. It is dimensionless (7). It usually refers to the first occurrence of the disease for each initially disease-free individual, although it is possible to consider the risk of developing the disease under observation within a specified period more than once (7).

In practice, risk is estimated by using different methods. The simple cumulative method is the easiest and most widely used (7). For a cohort of subjects followed for a given period of time, risk is often estimated by calculating the proportion of candidate subjects who develop the disease during the observation period. This measure is usually referred to as the cumulative incidence (CI) (7). One should be aware that in this case the term "cumulative incidence" plays a technical term, used more in a meaning of incidence proportion than in a meaning of incidence, cumulated over time. The observation period has to be clearly stated since the value of the measure is increasing with the prolongation of period of observation. This period could be based upon a calendar time, or not so (e.g. first year after the exposure, first year after surgery, etc.). Generally, cumulative incidence is estimated only for the first occurrence of the disease. If the durations of the individual follow-up periods for all non-cases are equal, the cumulative incidence is equivalent to the average risk for members of the cohort. This means that under the condition of a fixed cohort, cumulative incidence is a good estimate of risk. This is the reason that cumulative incidence and risk are frequently equalized. But once again, because risk is, by its definition, a conditional probability, it cannot be accurately estimated by calculating cumulative incidence unless all subjects in the observed candidate population are followed for the entire follow-up period or are known to develop the disease (or, another observed phenomenon) during the period of interest (7).

We have already presented the equation for calculation of incidence proportion (Equation 9), but since usually this measure is frequently denoted as risk (R), we repeat it as a new equation (Equation 12). In this module, this notation will be used from now on.

$$R = \frac{N_{d+\,new\,cases(gp)}}{N_{all\,persons\,at\,risk(bgp)}} \qquad \textbf{Equation 12.}$$

R = risk

$N_{d+\ new\ cases\ (gp)}$ = number of new cases of the disease under observation during a given period

$N_{all\ persons\ at\ risk\ (bgp)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given period

Calculation of this measure in practice is presented in "Case study 2".

Risk could be estimated using different methods (simple cumulative, actuarial, density, or Kaplan Meier method) (7,14,21), which will be discussed in a separate module in this book.

Incidence rate (in a classic epidemiologic sense) (1-3,23)

Under the term "incidence rate" many types of ratios are frequently referred, including proportions (21). One of them is the incidence rate in a classic epidemiologic sense. This measure is a ratio between new cases of the disease under observation in a given period of time and total number of the population at risk for getting a disease at the beginning of the given period, with suitable multiplier. Mostly, it is calculated by the following equation (Equation 13):

$$IR = \frac{N_{d+\,new\,cases(gp)}}{N_{all\,persons\,at\,risk(bgp)}} \times K \qquad \textbf{Equation 13.}$$

IR = incidence rate

$N_{d+\ new\ cases\ (gp)}$ = number of new cases of the disease under observation during a given period

$N_{all\ persons\ at\ risk\ (bgp)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given period

K = multiplier (1*00, 1000, 10.000, 100.000…)*

Calculation of this measure in practice is presented in Case study 2.

2. Measures in which the numerator is not included in the denominator

Incidence odds (1,2,22,25)

Incidence odds (also known as disease, or risk odds) is the measure of odds of getting ill during the period of observation. It is a ratio of conditional probability of developing the disease (risk) to conditional probability of not developing the disease (1-risk) (7).

If we abridge the elements of this ratio, odds represent the ratio of new cases of the disease under observation to persons who remained non-cases during the period of observation (Equation 14):

$$IO = \frac{R}{1-R} = \frac{N_{d+\,newcases(gp)}}{N_{d-\,(gp)}} \qquad \textbf{Equation 14.}$$

IO = incidence odds
$N_{d+\,newcases\,(gp)}$ = number of new cases of the disease under observation during a given period
$N_{d-\,(gp)}$ = number of all persons without the disease under observation (non-cases) during a given period

Calculation of this measure in practice is presented in "Case study 2".

This measure has no practical value, since in an incidence (cohort) study we can calculate an incidence proportion, or person-time incidence rate.

Incidence density (1,2,7,14,16,21,22)

Although, as it was emphasized previously, many types of ratios (including proportions) are frequently referred to as "rates", in its precise usage a "rate" is the ratio of a change in one quantity to a change in another quantity, with the denominator quantity often being time (21).

In measurement of incidence, there exist an index that measures how rapidly new cases of a phenomenon under observation are developing (when a death is a phenomenon under observation, how rapidly persons with a disease under observation are dying), or that measures the change in frequency of a health phenomenon to a change per unit of time. Some epidemiologists use the term "incidence rate" to denote this instantaneous measure (7,10), while others have referred to this concept as an instantaneous risk (7), or hazard rate (especially when death is the event under observation) (7,26,27). This index is measuring the instantaneous potential for change in disease status (from being disease-free to being diseased) per unit of time, relative to the size of the candidate population (disease-free population) at a given moment in time (1,7,27). If this measure is contrasted to incidence risk (incidence proportion), it is an instantaneous measure, which refers to a point in time and not to a period. Also, the incidence risk is dimensionless while person-time incidence rate is expressed in units of 1/time or $time^{-1}$ (e.g. $years^{-1}$) (7). In fact, it is the probability of the event under observation occurring within the time unit (e.g. day, month, year) under observation, given that it did not occur in a certain time unit (e.g. in a given day).

The problem of this measure is that we usually cannot express the size of the population at risk under observation (population free of disease at the beginning of the observation period) as a mathematical function of time, and thus we cannot express instantaneous incidence rates. Instead, we estimate an average incidence rate for a given period. This is analogous to the use of speed as an estimate of average velocity (7,10). The speedometer in a car is measuring how fast we are travelling at the moment of time we are looking at the speedometer. This does not mean that we are travelling with the same velocity all the time. The velocity is an example of an instantaneous rate. If we would read the velocity every few seconds for an hour, we could obtain an average velocity per hour. But there exist another measure, called speed that estimates the average velocity. The speed is the change in location divided by the change in time (we look at the kilometres counter at the beginning of the one-hour trip and at the end). The speed is an example of an average rate.

Coming back to epidemiologic data, it is much easier to determine an average rate than an instantaneous rate. Incidence density is an average rate for estimating average of instantaneous incidence rates (7,26). For this measure, other terms are used as well, being incidence rate, person-time incidence rate, average incidence rate, and force of morbidity (26)[1].

Technically, incidence density is the rate between the number of new cases which occur during the period under observation, and the quantity known under the term person-time (PT). This measure is expressed in units of 1/time or $time^{-1}$ as well. It could be calculated as follows (Equation 15):

[1] In this place, we need to give a warning about terminology used for this group of measures. A common problem in epidemiology of existence of multiple terms for the same concept is very explicit here. Also, a single term is applied to different concepts. The usage of terms differs from one author to another. On one hand, for example, Kleinbaum et al (7) are using the term "incidence rate" to denote the instantaneous potential per unit of time for event under observation to occur, given that the individual has survived up to the time (moment) of observation, and denote the average rate for estimating average of instantaneous rates as incidence density. The term "hazard rate" is used as a synonym for "incidence rate". On the other hand, for example, Benichou and Palta (26) are using the term "hazard rate" to denote the instantaneous potential per unit of time for event under observation to occur, given that the individual has survived up to the time (moment) of observation, and denote the average rate for estimating average of instantaneous rates as "incidence rate". The term "incidence density" is used as a synonym for "incidence rate".

$$ID = \frac{N_{d+newcases(gp)}}{PT}$$ **Equation 15.**

ID = incidence density
$N_{d+ new cases (gp)}$ = number of new cases of the disease under observation during a given period
PT = person-time

The measure is interpreted as average incidence rate for a cohort during the period under observation. Incidence density is the measure among family of incidence measures, which could play a role of an autonomous measure, or a role of intermediate measure in the process of estimating of incidence risk, what will be discussed in a separate module in this book.

As an element for calculation of incidence density, the quantity person-time or, person-time at risk, is introduced. It is the quantity which includes the information on number of individuals under observation at risk for getting the disease under observation (free of disease at the beginning of the observation period), and the exact time interval of this risk (the time between the beginning of the observation and the moment of break-out of the disease) (1). Mathematically, it is the sum of the periods of time at risk for each of the individuals under observation. This method enables to take into account how much of time exactly contributes each individual under observation to the population at risk, and thus to measure incidence rate over extended and variable time periods in a dynamic cohort in which there are several censored observations (deaths of other causes, change of domicile etc.). Usually, the time period is one year, and the measure is person-year (PY). In this concept each individual under observation contributes to the population at risk that many years as much as he/she was under observation before the disease under observation broke out (an individual under observation that is observed 1 year contributes 1 person-year, an individual under observation that is observed 9 months contributes 0.75 person-year, etc.). PY could be calculated as follows (Equation 16):

$$PY = t(y)_{ob1} + t(y)_{ob2} + \ldots + t(y)_{obn}$$ **Equation 16.**

PY= person-years
$t(y)_{ob1}$ = No. of years at risk for individual under observation No.1
$t(y)_{ob2}$ = No. of years at risk for individual under observation No.2
$t(y)_{obn}$ = No. of years at risk for individual under observation No.n

Calculation of PY and incidence density in practice is presented in “Case study 2”.

Special incidence measures

Mortality

Mortality is one of the most important epidemiologic and demographic measures which could be classified in the family of incidence measures (8,19). It is a ratio between number of deaths during a given period (usually 1 calendar year) and number of all persons at risk of dying during given period at the beginning of this period (usually number of the population, usually estimated at the middle of the year of the observation) (1,2,8). In fact, mortality is the incidence of death. Technically, it is usually expressed as rate (in a classic epidemiologic sense) (mortality rate, or death rate), and could be calculated as (Equation 17):

$$Mo = \frac{N_{deaths(gp)}}{N_{all\ persons\ at\ risk(bgp)}} \times K$$ **Equation 17.**

Mo = mortality rate
$N_{deaths (gp)}$ = number of deaths during a given period
$N_{all\ persons\ at\ risk\ (bgp)}$ = number of all persons at risk of dying during given period at the beginning of this period
K = multiplier

Hazard rate

We have already introduced the term »hazard rate« when we were introducing the concept of measuring instantaneous potential per unit of time for event under observation to occur, given that the individual has survived up to the time (moment) of observation. In the case when the observed event is a death from a disease, this measure is usually known as “hazard rate” (1,7,26,27). Similar to the idea of velocity, a hazard function h(t) gives the instantaneous potential at time t for getting an event. Estimate of average of these instantaneous potentials could be calculated as follows (27) (Equation 18):

$$h = \frac{N_{deaths(gp)}}{PT_d}$$ **Equation 18.**

h = hazard rate
$N_{deaths (gp)}$ = number of deaths during a given period
PT_d = person-time (sum of periods at risk for death for each individual)

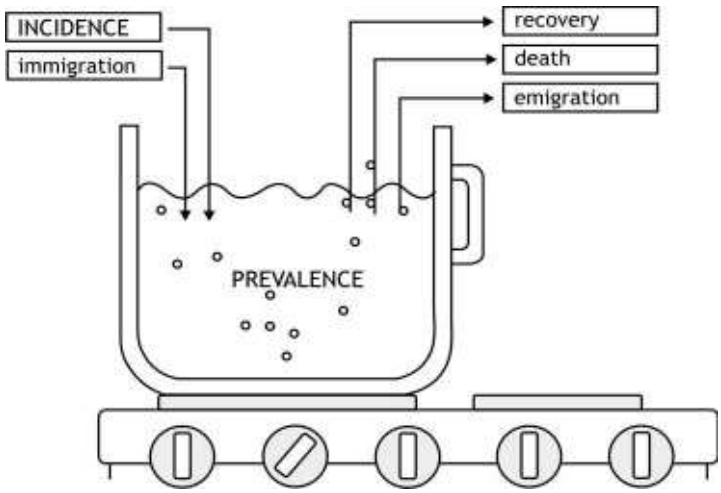
When the unit of time under observation is one day, this measure could be interpreted as the risk of dying for the person on that particular day, given that he/she has survived to that day (27).

Relationship between prevalence and incidence

Prevalence and incidence are very closely related (2,3).

This relationship is illustrated in Figure 3.

Figure 3. The relationship between prevalence and incidence of a health phenomenon



Input to the prevalence pool represent incident cases (new cases of a disease under observation), while output represent recoveries and deaths. Indirectly, the prevalence depends on duration of the disease. If the recovery rate is low, and the mortality is low as well, the chronicity of the disease is high and, consecutively, the prevalence is high. In such a case, even a low incidence leads to a high prevalence. Prevalence thus depends on incidence and duration of the disease. When both quantities are stable, and the prevalence of the disease is low (e.g. for cancer), this relationship could be expressed as follows (Equation 19) (2):

$$P = ID \times \text{average duration of the disease} \qquad \textbf{Equation 19.}$$

P = prevalence
ID = incidence density

However, incidence, recovery rate and mortality rate of the disease under observation are not the only factors which influence the prevalence. The smaller part of the prevalence pool input contribute also the immigrated cases, while the smaller part of the output contribute also emigrated cases. Furthermore, there exist the influence of competitive factors like mortality of extraneous factors (deaths because other causes than the disease under observation, e.g. traffic accidents). Because of a lot of possible influences, the prevalence has to be always interpreted cautiously.

Case studies

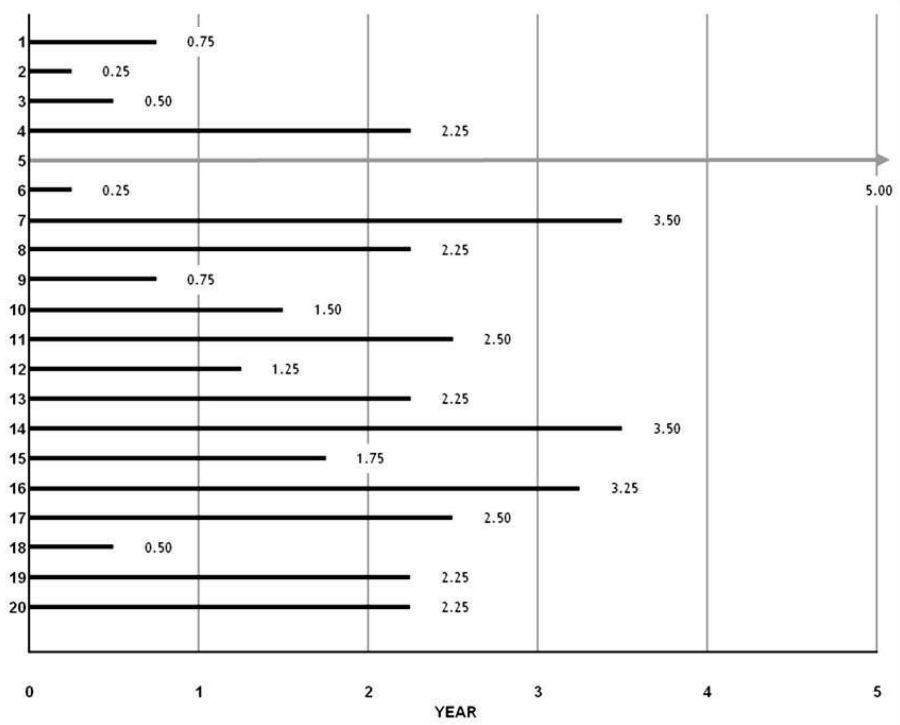
Data set

The illustration of differences between families of prevalence and incidence measures, and the differences between measures inside both families is be based upon an ideal set of example data (Figure 4).

The example data-set could be described as follows:

1. We have 20 individuals under observation which are all followed up for exactly 5 years. The course of events during 5-year observation time is shown in Figure 4. The most important example data characteristics are as follows:
   - at the beginning of the study all individuals under observation are without disease under observation, and
   - all of them are exposed to the effect of the same noxious agent,
   - some of them get ill and some not, and
   - all cases of disease are diagnosed.
2. Other important characteristics (for easier understanding of measures) are also:
   - all members enter the study at the same time (at the beginning of the study), and
   - nobody gets out the study (because of recovery, death, or change of domicile) – our cohort is a fixed cohort, all are followed-up exactly the same time,
   - the disease under observation is supposed to be chronic (there is no recovery after becoming diseased).

Figure 4. The course of events during a 5-year observation time of 20 individuals under observation, exposed to the effect of the same noxious agent, in the example data set



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Case study 1: Prevalence measures

Absolute and relative prevalence

For calculating the absolute prevalence let us choose the point 2 years after beginning of the study (Figure 5). Results of counting of existing cases of observed disease exactly 2 years after beginning of the study (Figure 5) show that there exist 9 persons with the disease (e.g. cases of the disease). Thus absolute prevalence or prevalent number, or simply prevalence, of the observed disease 2 years after beginning of the study is 9.

For calculating the relative prevalence let us choose again the point 2 years after beginning of the study (Figure 5). Results of counting of existing cases of observed disease exactly 2 years after beginning of the observation period (Figure 5) show that there exist 9 individuals under observation with the disease (e.g. cases of the disease) among a whole group of 20 individuals under observation. Relative prevalence as prevalence proportion or probability for having a disease under observation at point 2 years after beginning of the study is, when calculated according to Equation 3 as a decimal fraction, (Equation 20):

$$P_{rel\,(proportion)} = \frac{9}{20} = 0.450 \qquad \textbf{Equation 20.}$$

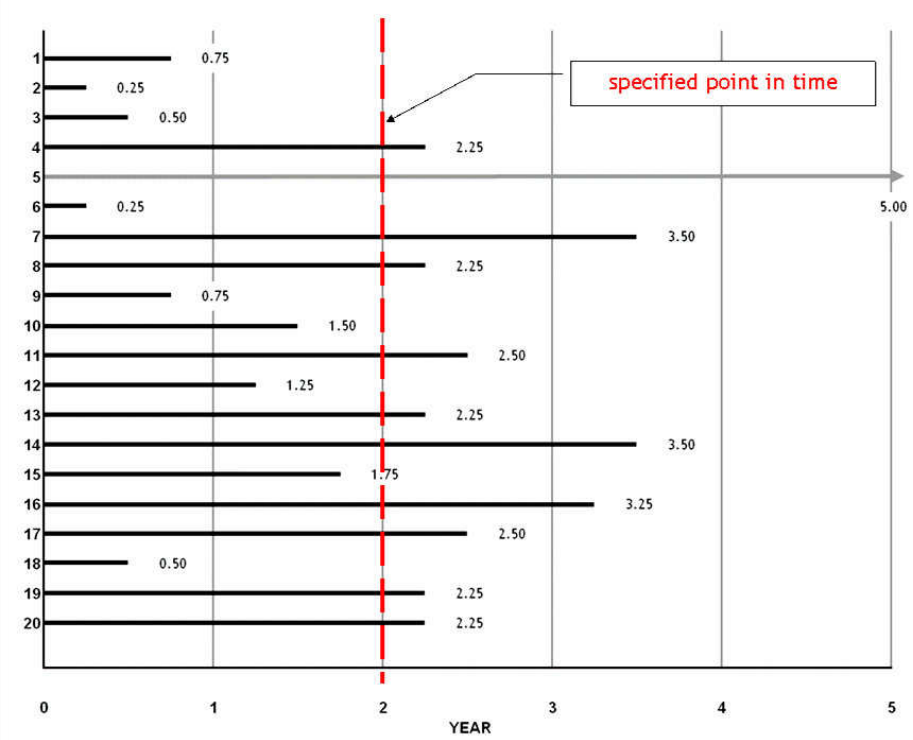
The relative prevalence expressed as prevalence proportion (probability of having a disease under observation) at point 2 years after beginning of the study is 0.450 or 45.0%.

Point and period prevalence

For calculating the point prevalence let us choose again the point 2 years after beginning of the study (Figure 5). The point prevalence in this point expressed as absolute point prevalence is 9, while expressed as a relative point prevalence (as a proportion) is according to Equation 4 (Equation 21):

$$_{point}P_{rel\,(proportion)} = \frac{9}{20} = 0.450 \qquad \textbf{Equation 21.}$$

Figure 5. Graphic presentation of point prevalence exactly two years after beginning of the observation period on example data



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

For calculating the period prevalence under assumption of fixed cohort, and assumption of chronicity of the disease under observation (once an individual gets the disease he/she cannot recover) let us choose the period of the second year of the study (Figure 6, dashed frame). The period prevalence in this period expressed as an absolute period prevalence is 9 - six individuals (No. 1, 2, 3, 6. 9 and 18) already had a disease at the beginning of the second year of the observation, while three of them got the disease during the second year period (No. 10, 12 and 15). The period prevalence, expressed as a relative period prevalence (as a proportion) is according to Equation 5 (Equation 22):

$$_{period}P_{rel(proportion)} = \frac{6+3}{20} = \frac{9}{20} = 0.450 \qquad \textbf{Equation 22.}$$

If the observed period is only the first half of the second year of the study (Figure 6 gray filled part of the dashed frame), the period prevalence in this period expressed as an absolute period prevalence is 7 - six individuals already had a disease at the beginning of the second year of the observation, while one of them got the disease during the first half of the second year period. The period prevalence, expressed as a relative period prevalence (as a prevalence proportion) is (Equation 23):

$$_{period}P_{rel(proportion)} = \frac{6+1}{20} = \frac{7}{20} = 0.350 \qquad \textbf{Equation 23.}$$

Relative point prevalence measures

Prevalence proportion

Results of counting of existing cases of observed disease exactly 2 years after beginning of the observation period (Figure 5) show that there exist 9 persons with the disease (e.g. cases of the disease) among 20. We have already calculated the prevalence proportion (Equations 19 and 20). We repeat this equation with a new notation according to Equation 6 (Equation 24):

$$P = \frac{9}{20} = 0.450 \qquad \textbf{Equation 24.}$$

Prevalence rate

Relative prevalence could be also expressed as prevalence rate (in a classic epidemiologic sense). In this case, it is calculated when the multiplier is 1,000 according to Equation 7 as follows (Equation 25):

$$PR = \frac{9}{20} \times 1{,}000 = 0.450 \times 1{,}000 = 450$$ **Equation 25.**

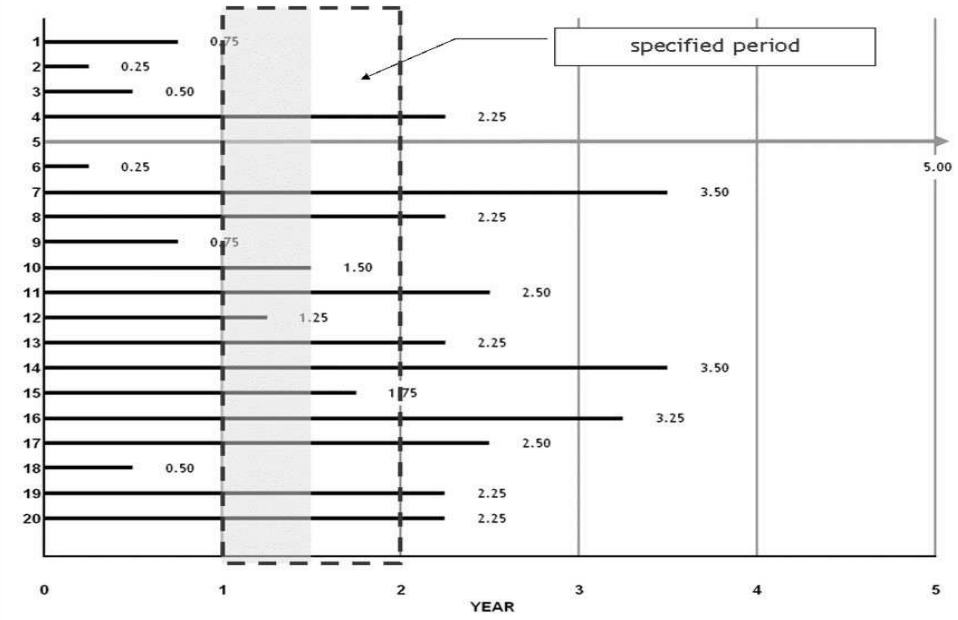
The relative prevalence of the disease under observation expressed as prevalence rate at point 2 years after beginning of the study is 450 per 1,000.

Prevalence odds

Results of counting of cases and non-cases of observed disease exactly 2 years after beginning of the observation period (Figure 5) show that there existed 9 persons with the disease (e.g. cases of the disease) and 11 without it (non-cases). Relative prevalence as prevalence odds for having a disease at this point in time is calculated according to Equation 8 as follows (Equation 26):

$$PO = \frac{9}{11} = 0.818$$ **Equation 26.**

Figure 6. Graphic presentation of period prevalence in the whole second year (dashed frame) or in the first half of the second year of the study (gray filled part of the dashed frame) on example data



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

The relative prevalence of the disease under observation expressed as prevalence odds at point 2 years after beginning of the study is 0.818. This means that 2 years after exposure there is 0.818 of a person with disease to 1 person without it (or if we calculate reverse odds – 1.222 of a person without a disease to 1 with it).

Case study 2: Incidence measures

Absolute and relative incidence

For calculating the absolute incidence let us choose the entire 5-year period of observation (Figure 7). Results of counting of cases of observed disease which broke out during the 5-year period (Figure 7) show that there were 19 cases, thus absolute incidence of the observed disease in a 5-year period of the study is 19.

For contrasting the relative incidence let us choose again the entire 5-year period of observation (Figure 7). Results of counting of cases of observed disease in which onset of this disease was registered during the 5-year time of observation (Figure 7) show that there were 19 cases among 20 individuals under observation. Relative incidence as incidence proportion during the 5-year period of observation is when calculated according to Equation 9 as a decimal fraction (Equation 27):

$$I_{rel\,(proportion)} = \frac{19}{20} = 0.950$$ **Equation 27.**

The relative incidence expressed as incidence proportion during the 5-year period of observation is 0.950 (or when expressed as a percentage, 95.0%).

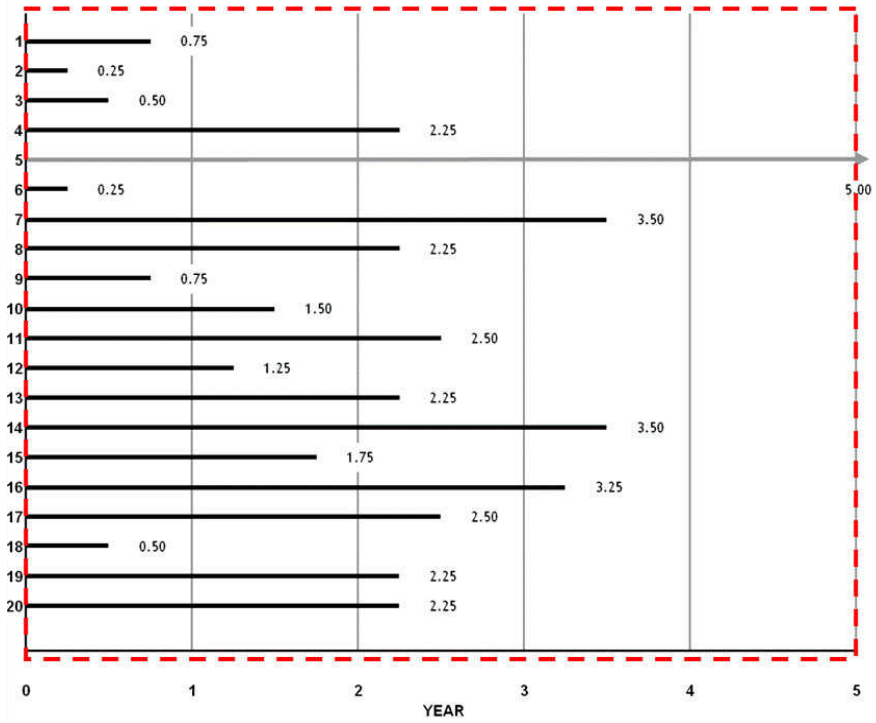
Cumulative and partial incidence proportion

For contrasting the cumulative and partial incidence let us choose again the entire 5-year period of observation (Figure 7). Results of counting of cases of observed disease which broke out during the entire 5-year time of observation (Figure 7) show that absolute cumulative incidence is 19, and relative cumulative incidence expressed as risk according to Equation 10 is as follows (Equation 28):

$$_{cum}I_{rel\,(proportion)} = \frac{19}{20} = 0.950 \qquad \textbf{Equation 28.}$$

Thus the cumulative incidence proportion for the 5-year period is 0.950 (or when expressed as a percentage, 95.0%).

Figure 7. Graphic presentation of incidence in the whole period of study on example data



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Results of counting of cases of observed disease which broke out during the first year of observation (Figure 8, dashed frame) show that there 6 cases of the disease appeared within this period thus absolute annual incidence in the first year of observation is 6. Relative annual incidence expressed as a proportion according to Equation 11 is as follows (Equation 29):

$$_{ann}I_{rel\,(proportion)(year\,1)} = \frac{6}{20} = 0.300 \qquad \textbf{Equation 29.}$$

Thus the annual incidence proportion for the first 1-year period is 0.300 (or when expressed as a percentage, 30.0%).

Relative annual incidences expressed as a proportion for the following four years are as follows (Equations 30-33):

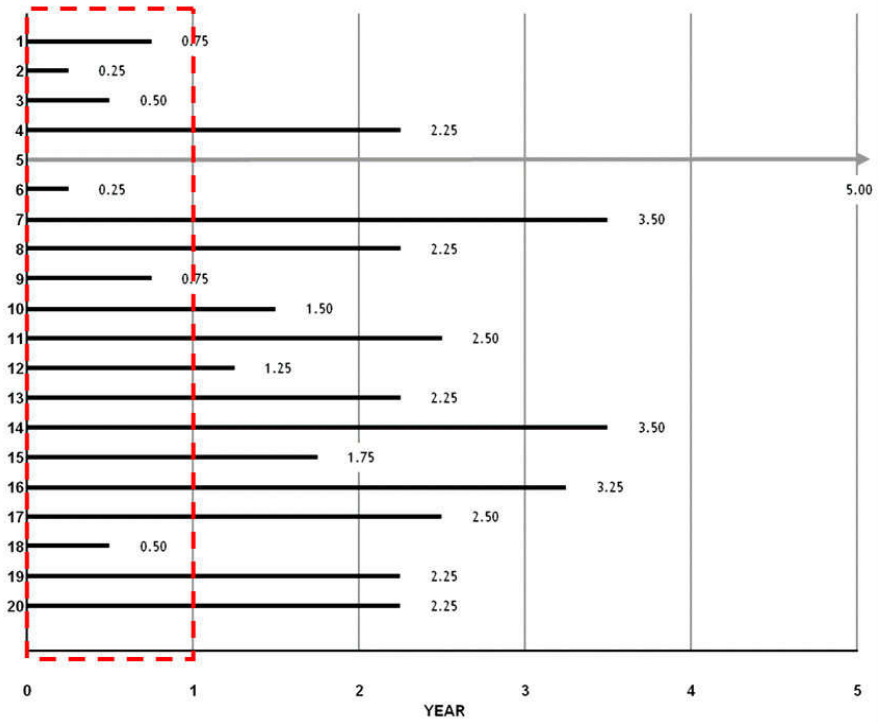
$$_{ann}I_{rel\,(proportion)(year\,2)} = \frac{3}{14} = 0.214 \qquad \textbf{Equation 30.}$$

$$_{ann}I_{rel\,(proportion)(year\,3)} = \frac{7}{11} = 0.636 \qquad \textbf{Equation 31.}$$

$$_{ann}I_{rel\,(proportion)(year\,4)} = \frac{3}{4} = 0.750 \qquad \textbf{Equation 32.}$$

$$_{ann}I_{rel\,(proportion)(year\,5)} = \frac{0}{1} = 0.000 \qquad \textbf{Equation 33.}$$

Figure 8. Graphic presentation of annual incidence in the first year of the study (red frame) on example data



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Relative incidence measures

Incidence risk (estimate)

Results of counting of cases of observed disease in which onset of this disease was registered during the 5-year time of observation (Figure 7) show that there were 19 cases among 20 individuals under observation. According to Equation 12, estimate of incidence risk is calculated as follows (Equation 34):

$$R = \frac{19}{20} = 0.950 \qquad \textbf{Equation 34.}$$

Thus, the incidence risk for the 5-year period of observation estimated by calculating cumulative 5-year incidence proportion is 0.950 (or when expressed as a percentage, 95.0%).

Incidence rate (in a classic epidemiologic sense)

Relative incidence could be also expressed as incidence rate (in a classic epidemiologic sense). In this case, it is calculated when the multiplier is 1,000 according to Equation 13 as follows (Equation 35):

$$IR = \frac{19}{20} \times 1{,}000 = 0.950 \times 1{,}000 = 950 \qquad \textbf{Equation 35.}$$

The relative incidence of the disease under observation expressed as incidence rate during the 5-year period of observation is 950 per 1,000.

Incidence odds

Results of counting of cases and non-cases of observed disease at the end of the 5-year period of observation show, that there exist 19 persons with the disease (e.g. cases of the disease) and 1 without it (non-case). Relative incidence as incidence odds of getting a disease during 5-year period according to Equation 14 is as follows (Equation 36):

$$IO = \frac{19}{1} = 19.000 \qquad \textbf{Equation 36.}$$

The relative incidence expressed as incidence odds at the end of the 5-year period of observation is 19.000. This means that in 5-year interval there will be 19 persons with disease to 1 person without it (or if we calculate reverse odds – 0.053 of a person without a disease to 1 with it).

Incidence density

The last relative incidence measure to be presented is incidence density. To calculate this measure we need first to calculate a quantity called person-years (Equation 16). Table 1 (based on Figure 4) is presenting the exact duration of time in which each individual under observation was under observation before the disease under observation broke out (time of being at risk). In total all 20 individuals under observation were exposed (at risk) in 5-year period 39 person-years.

Table 1. Data for calculation of person-years

Id. number	Onset of the disease (0=no, 1=yes)	Time of being at risk* (Years)
1	1	0.75
2	1	0.25
3	1	0.50
4	1	2.25
5	0	5.00
6	1	0.25
7	1	3.50
8	1	2.25
9	1	0.75
10	1	1.50
11	1	2.50
12	1	1.25
13	1	2.25
14	1	3.50
15	1	1.75
16	1	3.25
17	1	2.50
18	1	0.50
19	1	2.25
20	1	2.25
Total	Diseased: 19	Person-years: 39.00

In continuation, the incidence density for data presented in Figure 4 and Table 1 could be calculated for 5-year period according to Equation 15 as follows (Equation 37):

$$\mathrm{ID} = \frac{19}{39} = 0.4872 \qquad \textbf{Equation 37.}$$

If we than multiply the ID with 1,000 we get the value 487, which could be interpreted as: on average in 5-year interval 487 individuals under observation got ill per 1,000 population with the disease under observation if they are exposed to the effect of the noxious agent.

Exercises

Data set 1

A cohort of 20 individuals initially without a disease under observation, were followed up for 5 years. Times of events are presented in Figure 9.

Task 1

For the example set of data presented in Figure 9, please, calculate:

- absolute prevalence and relative prevalence as prevalence proportion at the point one years after beginning of the study.
- relative prevalence as prevalence proportion and prevalence odds at the point two years after beginning of the study.

Task 2

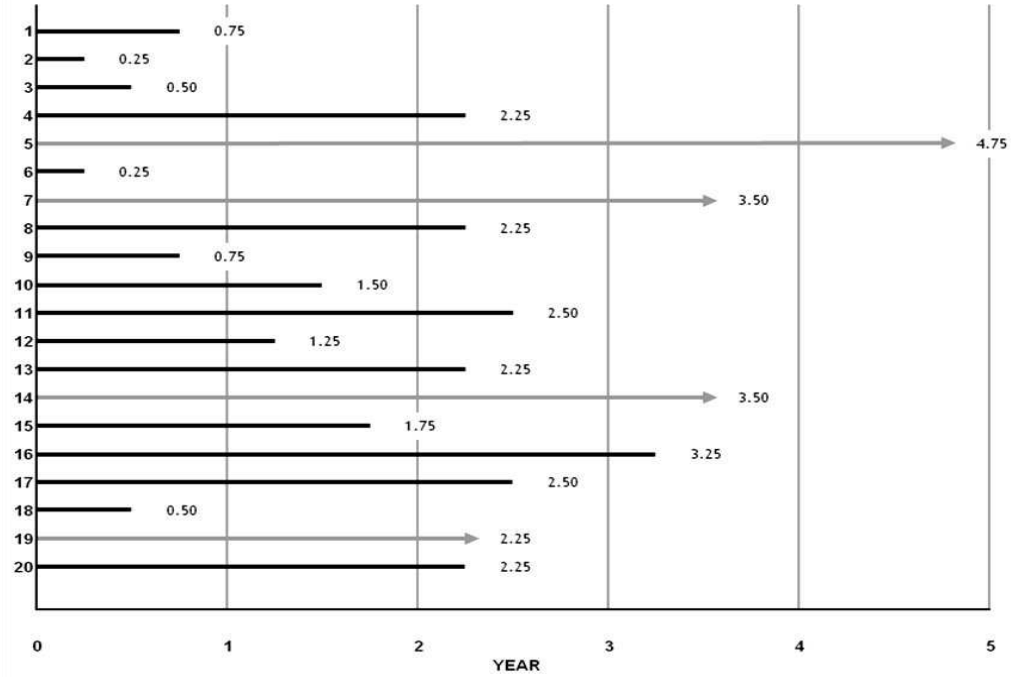
For the example set of data presented in Figure 9, please, calculate:

- cumulative incidence as a proportion for 5-year interval; discuss whether this measure is reliable or not,
- annual incidences as incidence proportion for each year of observation,
- incidence density for 5-year interval; discuss whether this measure is reliable or not.

Data set 2

In Figure 10, another imaginary data-set is presented. Again, a cohort of 20 individuals initially without a disease under observation, were followed up for 5 years.

Figure 9. Graphic presentation of events in a cohort of 20 people



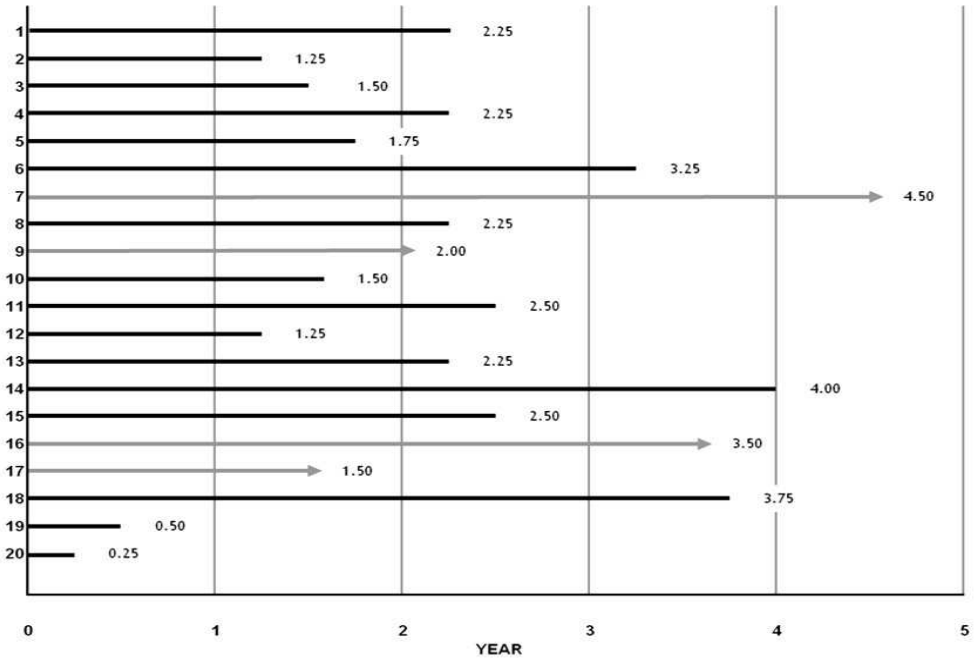
LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Task 3

For the example set of data presented in Figure 10, please, calculate:

- cumulative incidence as a proportion for 5-year interval; discuss whether this measure is reliable or not,
- annual incidences as incidence proportion for each year of observation,
- incidence density for 5-year interval; discuss whether this measure is reliable or not.

Figure 10. Graphic presentation of events in a cohort of 20 people



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

References

1 Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.

2. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999.
3. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
4. Anonymous. Prevalence and incidence. WHO Bull, 1966;35:783-784.
5. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization; 1998.
6. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Company; 1987.
7. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
8. Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. Fourth edition. London: BMJ Publishing Group; 1997. Available from URL: http://www.bmj.com/epidem/epid.html. Accessed: August 24, 2008.
9. Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005.
10. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.
11. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
12. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
13. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maidenhead: Open University Press, McGraw-Hill Education; 2005.
14. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
15. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
16. Centers for Disease Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics. Atlanta: Centers for Disease Control and Prevention. Available from URL: www.cdc.gov/osels/scientific_edu/ss1978/SS1978.pdf. Accessed September 5, 2013.
17. Babus V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
18. Fris RH, Sellers TA. Epidemiology for public health practice. Fourth edition. Sudbury, MA: Jones&Bartlett Publishers; 2009.
19. Kue Young T. Population health: concepts and methods. Second edition. Oxford: Oxford University Press; 2005.
20. Timmreck TC. An introduction to epidemiology. 3rd ed. Sudbury, MA: Jones&Bartlett Publishers; 2002.
21. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 5, 2013.
22. Pearce N. A short introduction to epidemiology. Second Edition. Occasional Report Series No 2. Wellington, NZ: Massey University Wellington Campus, Centre for Public Health Research; 2003.
23. Bayona M, Olsen C. Measures in epidemiology. Washington, D.C.: Young Epidemiology Scholars Program YES; 2004.
24. Kleinbaum DG, Klein M. Survival analysis. A self-learning text. Second edition. New York: Springer; 2005.
25. Pearce N. Effect measures in prevalence studies. Environ Health Perspect. 2004;112:1047-50.
26. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
27. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.

Recommended readings

1. Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005.
2. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization; 1998.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
5. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Age standardization: The direct method**
Module: 2.4	ECTS (suggested): 0.15
Author(s), degrees, institution(s)	Jadranka Bozikov, PhD, Professor Andrija Stampar School of Public Health, School of Medicine, University of Zagreb, Croatia Lijana Zaletel Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Doris Bardehle, MD, PhD, Professor Faculty of Health Sciences, University of Bielefeld, Germany
Address for correspondence	Jadranka Bozikov Andrija Stampar School of Public Health Rockefellerova 4, 10000 Zagreb, Croatia E-mail: jbozikov@snz.hr
Keywords	Age standardization, confounding, cumulative rate, direct standardization, indirect standardization, standardization.
Learning objectives	After completing this module students should: • understand the role of confounding in epidemiologic studies; • increase knowledge about methods of controlling of confounding in epidemiologic studies; • understand principles of direct standardization; • be capable to calculate age-standardized death rates using direct method.
Abstract	Basic theoretical background of standardization as one of the methods for controlling the effect of confounding in epidemiology is presented. Direct method of standardization as most common standardization method is presented in detail, using a case study. Step by step, the procedure is described using simple spreadsheet computer tool for facilitating it.
Teaching methods	Teaching methods include introductory lecture, exercises, and interactive methods such as small group discussions. After introductory lectures, students first carefully read the recommended sources in age standardization. Afterwards, they discuss with other students the standardization as a method of controlling confounding. In continuation, in groups of 2-3 students they perform the procedure of direct standardization using the programme tool (e.g. MS Excel) on given data. At the end, students compare and discuss their results.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (1 computer per 2-3 students), LCD projection, access to the Internet; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaire (MCQ) and case-study.

AGE STANDARDIZATION: THE DIRECT METHOD

Jadranka Bozikov, Lijana Zaletel-Kragelj, Doris Bardehle

Theoretical background

Population diversity and confounding

When examining the health of populations, one of the fundamentals of this process is the comparison of health indicators among and/or across different population subgroups within the countries.

Whenever we want to compare epidemiologic measures, irrespective of what they represent: morbidity (e.g. incidence or prevalence measures), mortality or other measure, across different populations or population groups we should take into account their diversity (1). Namely, populations/population groups are heterogeneous in regard to various health related characteristics (e.g. age, gender, education, religion, genetic and geographic factors, etc.) (2).

When the epidemiological measures are calculated without taking into account this diversity, such kind of epidemiological measures are called crude measures. The potential influence of the diversity could be imagined if the procedure of calculation of crude values is taken into consideration - the value of crude population measure is in fact an average of the values for the individual subgroups within a population (e.g. subgroups according to age), weighted by their relative sizes (1). This means that, the larger the subgroup (e.g. age subgroup), the more influence it will have on the crude measure. The comparison of crude measures across populations (or population groups) can be thus misleading because they can be greatly affected by the influence of such characteristics (e.g. different age distributions in the populations/population groups being compared).

In statistical terms, these characteristics are referred to as confounders. Confounding (from the Latin "confundere" that means to mix together) is according to Last et al. defined as an effect which appears when the measurement of the effect of an exposure on a risk is distorted by the relation between the exposure and other "extraneous" factor (or multiple factors) that also influence the outcome under study (3). In this context, extraneous factors are considered as factors other than the relationships between two phenomena under study. But, not every characteristic meets the criteria for being a confounder. A confounding factor (or confounder) must meet three criteria:

- it should be a known risk factor for the result of interest (4),
- it should be a factor associated with the exposure, but not a result of exposure (4), and
- it should be a factor that is not an intermediate variable between them.

Thus, when crude rates are interpreted, this interpretation would have been confounded by differences in the populations being compared (e.g. differences in age distribution). We, therefore, need to control for the effects of confounders in order to remove the confounding effect.

Controlling for the effects of confounding

There exist several procedures to control the effects of confounding. Some of them could be performed in the designing and planning phase of a study, and the others in the phase of data analysis (5-7). The first group of procedures (e.g. randomization, restriction, matching) is usually performed in experimental studies while the second group (stratification, standardization, statistical modelling) is conducted in observational studies (5-7). This concept of control of confounding in epidemiology derives from the limited opportunities for experimental control in non-experimental design of studies.

In practice, age is the factor that is most frequently controlled or adjusted for confounding. In an older population, higher rates of certain diseases that more frequently appear in older age-groups (e.g. cancers) could be observed not because of the presence of risk factors, but because of the higher age itself (8). Traditionally, in controlling for age confounding, standardization is applied (8).

Standardization

Definition and description

Standardization of health indicators is a classic epidemiological method defined as:

- a set of techniques used to remove as much as possible the effects of differences in age or other confounding variables when comparing two or more populations (3),
- a method that removes the confounding effect of variables that we know (or think) that they could influence the comparison between two or more populations (5,6),
- a statistical method for deriving measures that are comparable across populations that differ in age and other demographic characteristics (9).

Standardization provides an easy-to-calculate and easy-to-use summary measure e.g. standardized mortality (abbreviated sometimes as SMR[2]) or standardized death rate (abbreviated as SDR[3]) when the outcome is death, or a standardized morbidity measure when the outcome is disease occurrence (e.g. standardized incidence rate in the case the morbidity measure is incidence - abbreviated sometimes as SIR[4]). These measures can be useful for information users, such as decision-makers.

Types of standardization

Two approaches to standardization could be used, direct and indirect (1,3,5-9). They are used in different situations that will be described below.

Direct standardization

Direct standardization is a procedure that forms a weighted average of age specific rates or risks, using as weights the distribution of a specified standard population (1,3,8,9).

The method is called "direct" because it uses the actual morbidity or mortality rates of the populations being compared (9).

In the direct standardization method, according to Last et al. (3), the directly standardized rate represents what the crude rate would have been in the observed population if that population had the same structure as the standard population with respect to the variable (or more variables) for which the standardization was performed.

Thus, these rates are hypothetical and by themselves they are not meaningful because they are not real. These rates are useful only if they are used in comparisons of populations in the case that standardized rates in all the compared populations are derived by the same procedure using the same standard population.

Direct standardization could be used to compare observed populations for which the specific crude rates are known and statistically stable. It is commonly used in reports of vital statistics (e.g., mortality) or major disease incidence trends (e.g., cancer incidence).

Indirect standardization

Indirect standardization is used to compare observed populations for which the specific crude rates are unknown, or are statistically unstable (3). This is frequently the case of small populations, or when the observed phenomenon is rare.

It is different from direct standardization in both, the method and interpretation. Instead of using the structure of the standard population, we utilize its specific rates and apply them to the populations under comparison, previously stratified by the variable to be controlled for. The total of expected cases is obtained this way. The SDR is then calculated by dividing the total of observed cases by the total of expected cases. This ratio allows comparison of each population under study to the standard population. A conclusion can be reached by simply calculating and looking at the SDR. A SDR higher than one (or, 100% if expressed in percentage) indicates that the risk of dying in the observed population is higher than what would be expected if it had the same experience or risk as the standard population. On the other hand, a SDR lower than one (or, 100%) indicates that the risk of dying is lower in the observed population than expected if its distribution were the same as the reference population.

Indirect standardization plays a major role in studies of occupational diseases.

Age standardization

Although age standardization is not a special type of standardization, we think it is worthy to emphasize it. As already mentioned, age is the factor that is most frequently standardized for, since the age is one of the most important confounders. Compared populations could have very different age structure that can influence the interpretation of differences in crude rates of observed phenomenon.

Age-standardized rates calculated using the direct method represent what the crude rate would have been if the population had the same age distribution as the standard population.

Age-standardization is particularly used in comparative mortality studies, since the age structure has an important impact on a population's overall mortality.

Limitations of standardization

It is important to know that standardization as a method for controlling confounding has some limitations. Any summary measure can hide patterns that might have important public health implications. For example, with age standardization, one might fail to detect age-specific differences in risk across time or place. This might arise if a disease is displaying an increasing incidence due to a birth cohort effect (people at younger ages might have a higher risk in recent years compared to previous years, while older people could have the opposite pattern). An age-

[2] We should be aware that this abbreviation, SMR, is also used in the case of standardized mortality ratio as an outcome measure in indirect standardization procedure which is not the subject of this module.

[3] The term standardized death rate (abbreviated as SDR) is commonly used in Health for All Data Base of WHO, European Region (10).

[4] The same as under 1.

standardized rate could hide these trends. Despite this risk, standardized rates have proved to be very useful summary measures.

The procedure of direct age standardization

Entry data for the procedure

For accomplishing the procedure of direct age standardization, we need three sets of data:

1. Number of cases of a health phenomenon (death, disease) to be standardized:
   We need absolute frequency (number of cases) of a health phenomenon to be standardized across the age groups.
   These data are usually derived from registration of health phenomena (mortality, morbidity data) - health statistics of a country. Usually these data are administered by national public health institutes. Mortality data are usually available, while morbidity data (e.g. cancer incidence) are more difficult to obtain. In Slovenia for example, cancer incidence for several sites could be obtained from a high quality Cancer Registry of the Republic of Slovenia. The Registry's annual reports, Cancer Incidence in Slovenia, are one of the regular ways of disseminating information of this registry. They are publicly available from their homepage as PDF files (11).
2. Observed population data:
   The next set of data that is needed for direct standardization is the distribution of population according to age.
   These data are usually derived from the on-going registration of populations and/or population censuses. They are usually provided by every country's statistical office. For example, for Slovenia these data are provided by the Statistical Office of the Republic of Slovenia. They are publicly available in Office's annual reports, Statistical Yearbook, from their homepage as PDF files (12).
   For most of countries worldwide, these data can also be obtained from the U.S. Census Bureau International Data Base Entry (13).

Table 1. Some standard populations. Adapted from Health for All database Manual (15) and Surveillance, Epidemiology and End Results (SEER) Program homepage (16)

Age group	European standard population (100,000)	European standard population (million)	World standard population (million)	1996 Canadian standard population (million)	2000 US standard population (million)
0	1,600	16,000	24,000	12,342	13,818
1-4	6,400	64,000	96,000	53,893	55,317
5-9	7,000	70,000	100,000	67,985	72,533
10-14	7,000	70,000	90,000	67,716	73,032
15-19	7,000	70,000	90,000	67,841	72,169
20-24	7,000	70,000	80,000	67,761	66,478
25-29	7,000	70,000	80,000	72,914	64,529
30-34	7,000	70,000	60,000	87,030	71,044
35-39	7,000	70,000	60,000	88,510	80,762
40-44	7,000	70,000	60,000	80,055	81,851
45-49	7,000	70,000	60,000	71,847	72,118
50-54	7,000	70,000	50,000	55,812	62,716
55-59	6,000	60,000	40,000	44,869	48,454
60-64	5,000	50,000	40,000	40,705	38,793
65-69	4,000	40,000	30,000	37,858	34,264
70-74	3,000	30,000	20,000	32,589	31,773
75-79	2,000	20,000	10,000	23,232	26,999
80-84	1,000	10,000	5,000	15,424	17,842
85+	1,000	10,000	5,000	11,617	15,508
Total	100,000	1,000,000	1,000,000	1,000,000	1,000,000

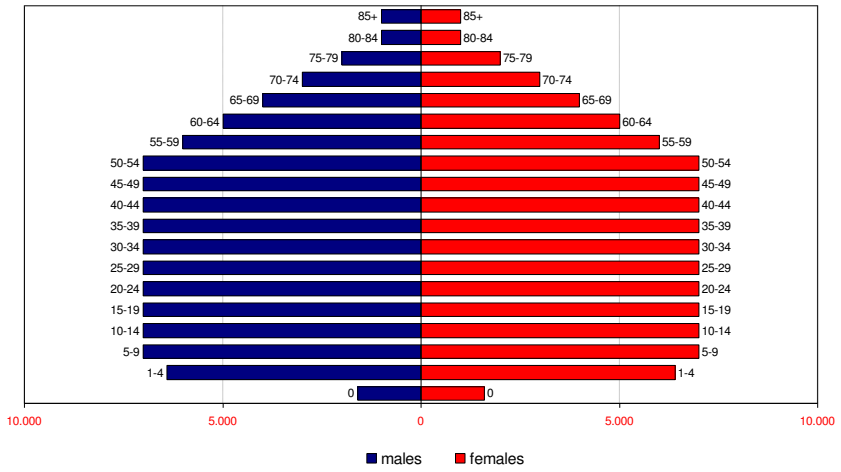
3. Standard population data:
   An important step in the direct standardization is the selection of a standard population (4), since the value of the adjusted rate depends on the standard population used.
   The standard population may come from the populations under study – average, for example. In this case, however, it is important to ensure that the populations do not differ considerably in their size, since a larger

population may influence the adjusted rates (14). The standard population may also be a population without any relation to the data under study but, in general, its distribution with regard to the adjustment factor should not be radically different from the populations we want to compare.

In the European region of the World Health Organization, for comparison across countries within this region, age-standardized death rates are calculated using the European standard population, while in other regions there are employed other standard populations. The detailed description of the European standard population could be obtained from the European Health for All Database manual (15). The age distribution of four different hypothetical standard populations is presented in Table 1 (15,16).

The European standard population, which will be used in our case study, is also presented in Figure 1. For all three sets of entry data the same age distribution is needed.

Figure 1. European standard population (100,000). Adapted from the Health for All database Manual (15)



The procedure

Directly standardized rate is, in general, calculated by dividing the number of deaths by the actual local population in a particular age group multiplied by the standard population for that particular age group and summing across the relevant age groups. The rate is usually expressed per 100,000. The exact procedure for calculating standardized death rates in 4 steps is as follows:

1. Step 1 - calculation of the specific crude death rate for every (specific) age group.
   The crude specific death rate for every age group is obtained by dividing the number of deaths in every specific age group by the observed (actual local) population in this age group multiplied by a multiplier (usually 100,000) (Equation 1):

$$_{crude}DR_{(spec.group)} = \frac{N_{deaths(spec.group)}}{N_{pop(spec.group)}} \times 100{,}000 \qquad \textbf{Equation 1.}$$

   $_{crude}DR_{(spec.\ group)}$ = crude death rate in a specific population group
   $N_{deaths(spec.\ group)}$ = number of deaths in a specific population group
   $N_{pop(spec.\ group)}$ = number of population in a specific population group

2. Step 2 - calculation of the crude rate for total population.
   The crude rate for total population is calculated using the similar formula as in calculating specific death rate for every age group (Equation 1), except that in this calculation totals of number of cases and population are used (Equation 2).

$$_{crude}DR_{(total\ pop)} = \frac{N_{deaths(total\ pop)}}{N_{pop(total\ pop)}} \times 100{,}000 \qquad \textbf{Equation 2.}$$

   $_{crude}DR_{(total\ pop)}$ = crude death rate in a total population
   $N_{deaths\ (total\ pop)}$ = number of deaths in a total population
   $N_{pop\ (total\ pop)}$ = number of population in a total population

   These totals need to be calculated prior to calculation of the crude rate for total population.

3. Step 3 - calculation of the expected number of deaths in the standard population for every specific age group.
   The expected number of deaths in a specific age group is calculated by multiplying the result obtained in step 1 by the number of population in the standard population in this specific age group and dividing it by the multiplier used in step 1 (usually 100,000) (Equation 3):

$$N_{exp.deaths(spec.group)} = \frac{{}_{crude}DR_{(spec.group)} \times N_{stand.pop(spec.group)}}{100{,}000} \qquad \textbf{Equation 3.}$$

$N_{exp.deaths\ (spec.\ group)}$ = number of expected deaths in the standard population in a specific population group

${}_{crude}DR_{(spec.\ group)}$ = crude death rate in a specific population group

$N_{stand.\ pop(spec.\ group)}$ = number of population in a specific population group of a standard population

The result of this step, the expected number of deaths in every specific age group, is in fact the standardized death rate in this particular age group.

4. Step 4 - calculation of the standardized death rate in the total population.
   Finally, the standardized death rate is obtained by summation of the expected number of deaths in a specific age group across all age groups (Equation 4).

$$ {}_{stand}DR_{(total\ pop)} = \sum N_{exp.deaths(spec.group)} \qquad \textbf{Equation 4.}$$

${}_{stand}DR_{(total\ pop)}$ = standardized death rate in a total population

$N_{exp.\ deaths\ (spec.\ group)}$ = number of expected deaths in the standard population in a specific population group.

Case study

The procedure of direct age standardization of disease D mortality in Croatia

Entry data

For accomplishing the procedure of direct age standardization of disease D mortality in Croatia we need the following sets of data:

1. Number of deaths of a disease D to be standardized:
   In table 2, the number of cases of the disease D in every age group for the male population for year 2000 is presented. The data were obtained from National Health Institute of Croatia (17).
2. Observed population data:
   In Table 3, the number of population in every age group of the Croatian population is presented. The 1991 census data are used (Table 3, Figure 2). The data were obtained from Central Bureau of Statistics of Republic of Croatia (18). The same data are presented also in Figure 2.
3. Standard population data:
   Given the fact that we want to compare the mortality rate of disease D in Croatia to other countries of World Health Organization (WHO) European region mortality rates, the European standard population is the best choice for standard population. This standard population has been already presented (Table 1, Figure 1).

Table 2. Number of death cases (absolute incidence) of the disease D in Croatia for male population for every age group for year 2000. Source: National Health Institute of Croatia (17)

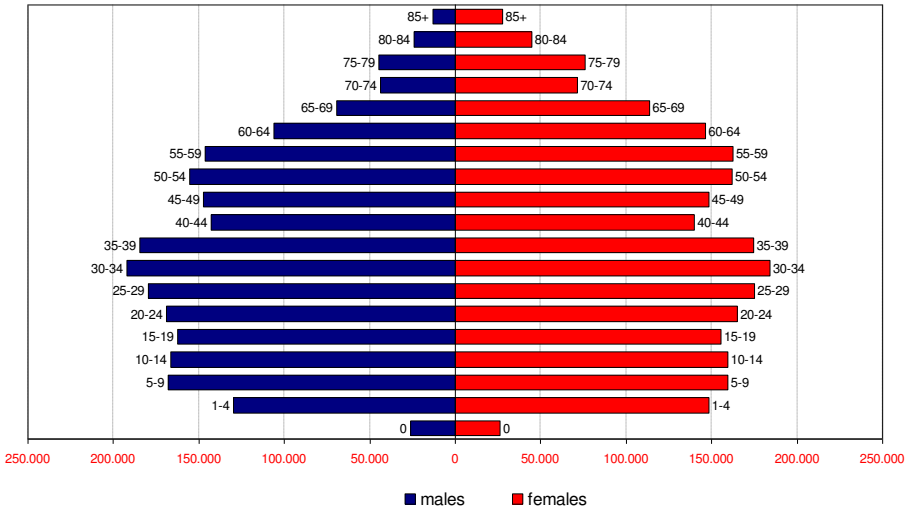
Age group	Number of cases	Age group	Number of cases
0	0	45-49	80
1-4	0	50-54	143
5-9	0	55-59	237
10-14	2	60-64	258
15-19	3	65-69	249
20-24	4	70-74	200
25-29	15	75-79	253
30-34	14	80-84	159
35-39	36	85+	68
40-44	52	Total	1773

Table 3. Croatian population (census 1991) in figures

Age group	Males	Females	Age group	Males	Females
0	26361	26361	45-49	147304	148308
1-4	130000	148272	50-54	155474	161793
5-9	168031	159688	55-59	146177	162304
10-14	166573	159218	60-64	105909	146527
15-19	162383	155564	65-69	69655	113449
20-24	169107	164779	70-74	43815	71653
25-29	179330	175245	75-79	44536	75999
30-34	192397	184039	80-84	23986	44564
35-39	184654	174497	85+	12844	27651
40-44	142937	139918	Total	2271473	2439829

Source: Central Bureau of Statistics of Republic of Croatia (18)

Figure 2. Number of population by sex in nineteen age groups of the Croatian population, the 1991 census data. Source: Central Bureau of Statistics of Republic of Croatia (18)



Setting the frame table for the standardization procedure

The procedure for standardization of rates could be automatised by using an appropriate computer programmes. Spreadsheet programme like Microsoft Excel could be used. The frame table for the procedure should contain the following columns:

- Age group,
- Number of cases of deaths,
- Population (observed),
- Rate per 100,000,
- European standard population, and
- Expected cases of deaths (in European standard population).

In Figure 3, this frame is presented, while in Figure 4, in this frame entry, the data are already filled in.

Figure 3. The frame table for the standardization procedure in Microsoft Excel computer programme

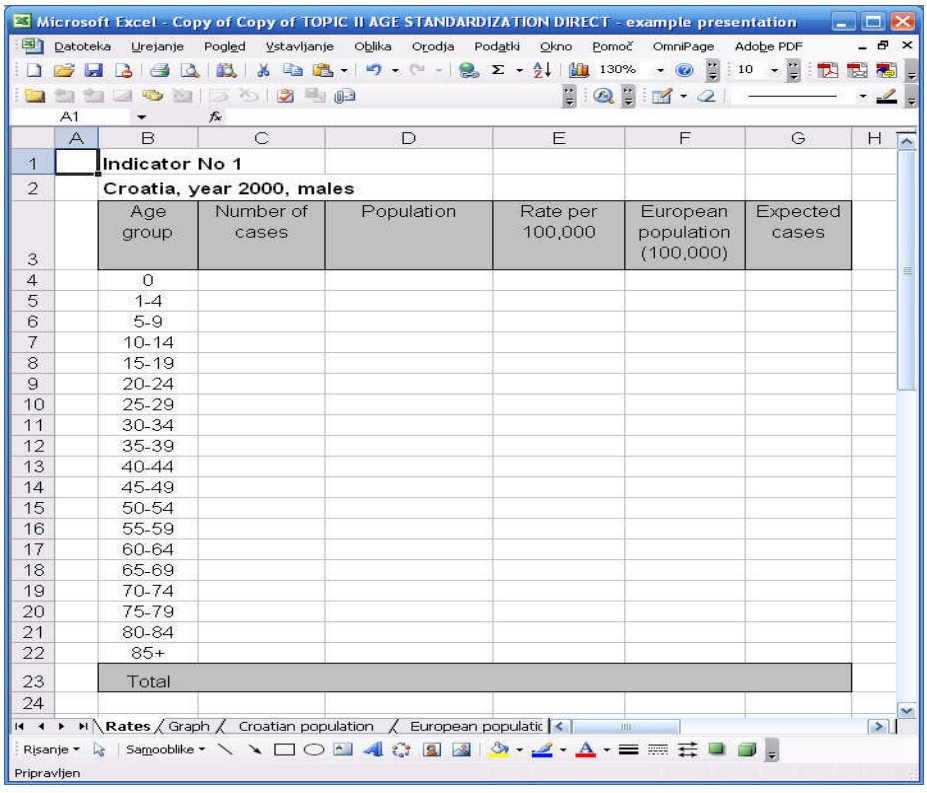
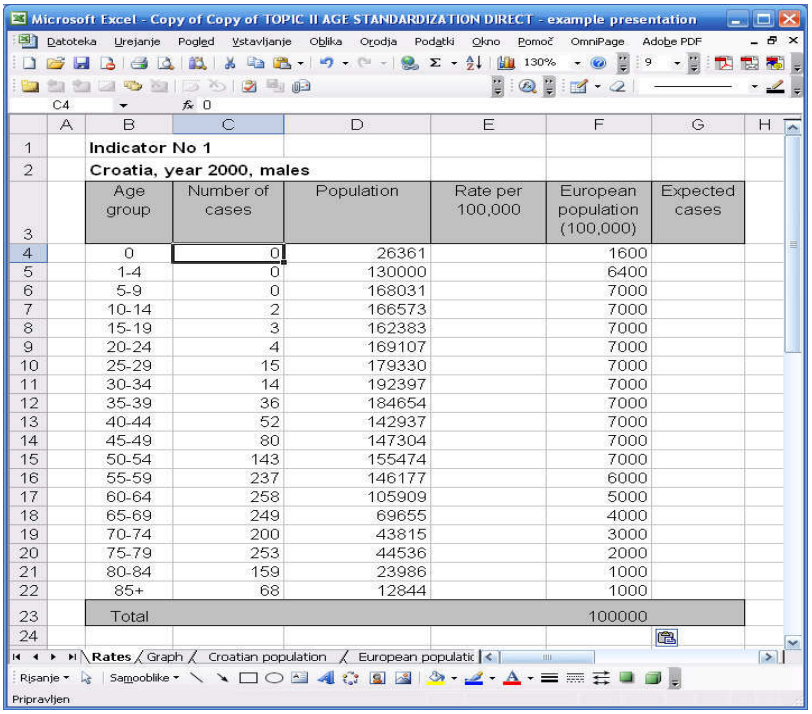


Figure 4. The frame table for the standardization procedure in Microsoft Excel computer programme filled in with entry data



Microsoft Excel - Copy of Copy of TOPIC II AGE STANDARDIZATION DIRECT - example presentation

	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361		1600		
5		1-4	0	130000		6400		
6		5-9	0	168031		7000		
7		10-14	2	166573		7000		
8		15-19	3	162383		7000		
9		20-24	4	169107		7000		
10		25-29	15	179330		7000		
11		30-34	14	192397		7000		
12		35-39	36	184654		7000		
13		40-44	52	142937		7000		
14		45-49	80	147304		7000		
15		50-54	143	155474		7000		
16		55-59	237	146177		6000		
17		60-64	258	105909		5000		
18		65-69	249	69655		4000		
19		70-74	200	43815		3000		
20		75-79	253	44536		2000		
21		80-84	159	23986		1000		
22		85+	68	12844		1000		
23		Total				100000		
24								

Rates / Graph / Croatian population / European populatic

Risanje | Samooblike

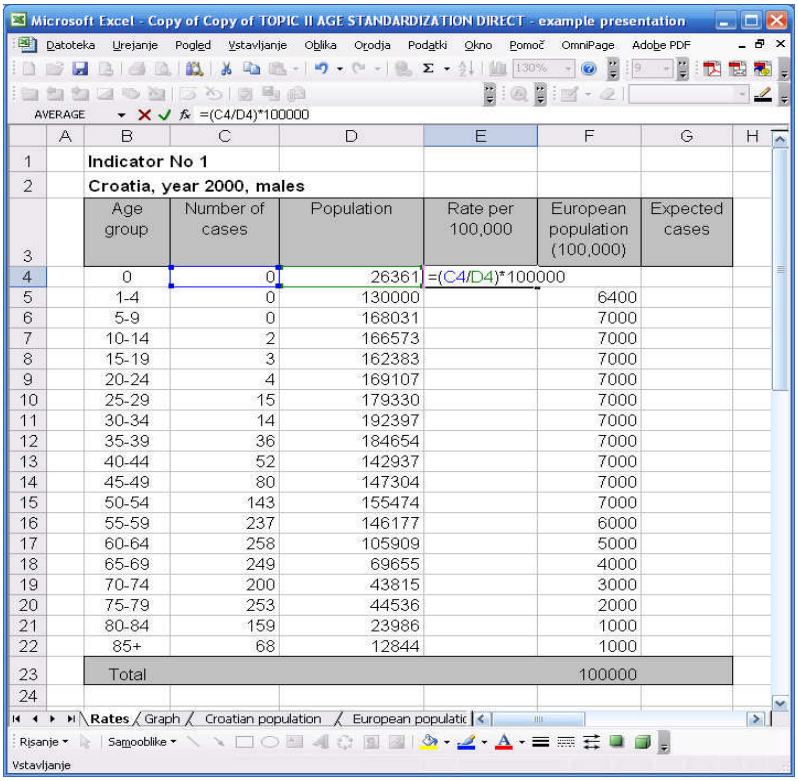
Pripravljen

The procedure

The four steps are as follows:

1. Step 1 - calculation of the specific crude death rate for every (specific) age group.
   The crude specific death rate for every age group is calculated by using the Equation 1. In Figure 5 is presented the equation for calculating the crude specific death rate for the age group 0 using corresponding cells for number of deaths and the observed (actual local) population in this age group in a spreadsheet.

Figure 5. Calculation of the specific rate for age group 0 using corresponding cells for number of deaths, and the observed (actual local) population in this age group in a spreadsheet



Microsoft Excel - Copy of Copy of TOPIC II AGE STANDARDIZATION DIRECT - example presentation

AVERAGE | =(C4/D4)*100000

	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	=(C4/D4)*100000			
5		1-4	0	130000		6400		
6		5-9	0	168031		7000		
7		10-14	2	166573		7000		
8		15-19	3	162383		7000		
9		20-24	4	169107		7000		
10		25-29	15	179330		7000		
11		30-34	14	192397		7000		
12		35-39	36	184654		7000		
13		40-44	52	142937		7000		
14		45-49	80	147304		7000		
15		50-54	143	155474		7000		
16		55-59	237	146177		6000		
17		60-64	258	105909		5000		
18		65-69	249	69655		4000		
19		70-74	200	43815		3000		
20		75-79	253	44536		2000		
21		80-84	159	23986		1000		
22		85+	68	12844		1000		
23		Total				100000		
24								

Rates / Graph / Croatian population / European populatic

Risanje | Samooblike

Vstavljanje

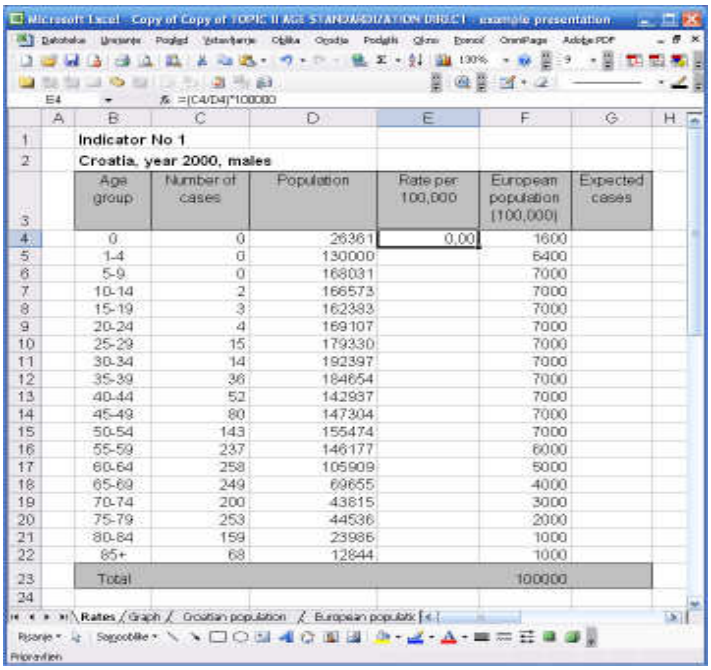
In Figure 6, the results of this step is presented.

In Equation 5, the procedure for calculating the crude specific death rate for the age group 40-44 is presented, as well as the result.

$$_{\text{crude}}\text{DR}_{\text{age40-44}} = \frac{52}{142{,}937} \times 100{,}000 = 36.38 \qquad \textbf{Equation 5.}$$

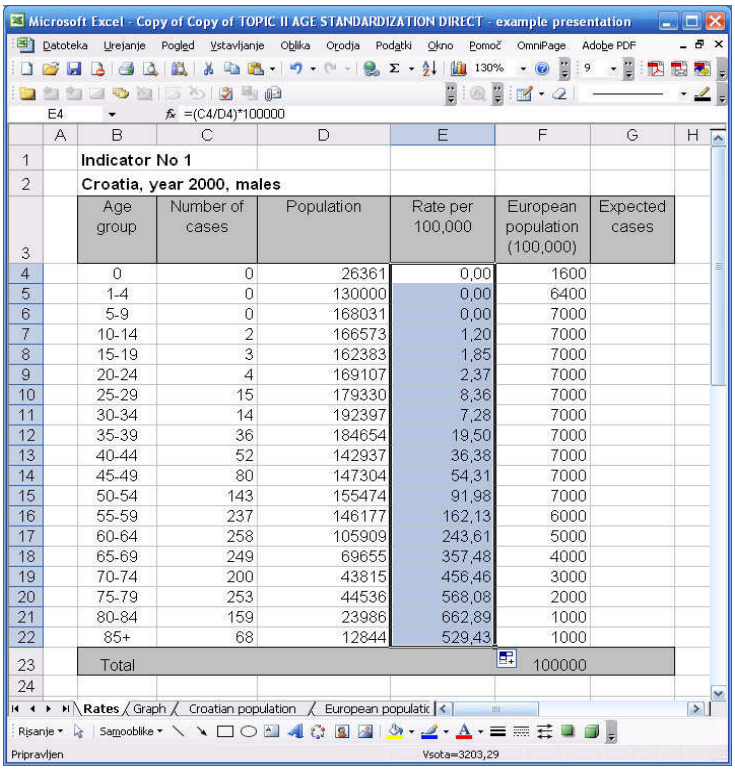
In continuation, the formula from the cell containing the function for calculating the crude specific death rate for the age group 0 is copied to other cells in the same column by dragging the right lower corner of the cell and extending it to the last age group cell (Figure 7).

Figure 6. The result of calculation of the specific rate for age group 0



	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	0,00	1600		
5		1-4	0	130000		6400		
6		5-9	0	168031		7000		
7		10-14	2	166573		7000		
8		15-19	3	162383		7000		
9		20-24	4	169107		7000		
10		25-29	15	179330		7000		
11		30-34	14	192397		7000		
12		35-39	36	184654		7000		
13		40-44	52	142937		7000		
14		45-49	80	147304		7000		
15		50-54	143	155474		7000		
16		55-59	237	146177		6000		
17		60-64	258	105909		5000		
18		65-69	249	69655		4000		
19		70-74	200	43815		3000		
20		75-79	253	44536		2000		
21		80-84	159	23986		1000		
22		85+	68	12844		1000		
23		Total				100000		
24								

Figure 7. Copying of the function used for calculation of the specific rate for age group 0 to all age groups by dragging the right lower corner of the cell and extending it to the last age group cell

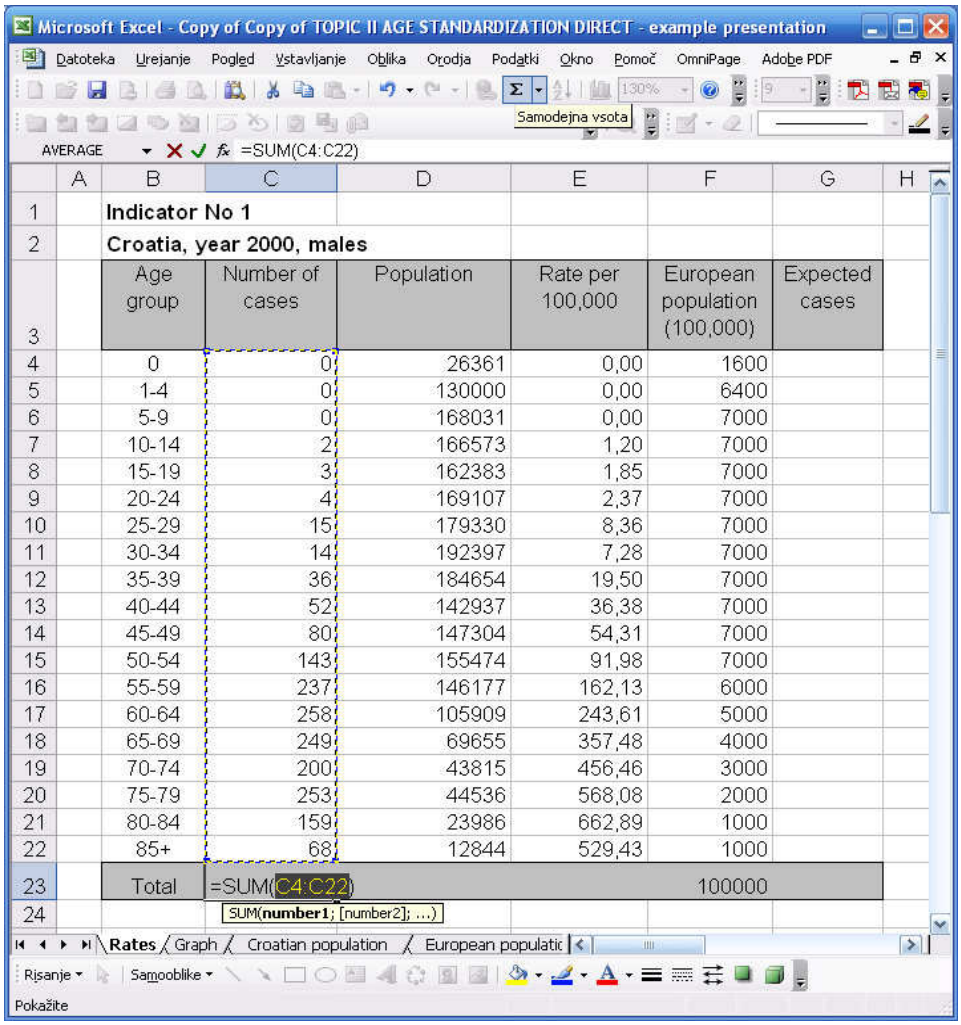


	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	0,00	1600		
5		1-4	0	130000	0,00	6400		
6		5-9	0	168031	0,00	7000		
7		10-14	2	166573	1,20	7000		
8		15-19	3	162383	1,85	7000		
9		20-24	4	169107	2,37	7000		
10		25-29	15	179330	8,36	7000		
11		30-34	14	192397	7,28	7000		
12		35-39	36	184654	19,50	7000		
13		40-44	52	142937	36,38	7000		
14		45-49	80	147304	54,31	7000		
15		50-54	143	155474	91,98	7000		
16		55-59	237	146177	162,13	6000		
17		60-64	258	105909	243,61	5000		
18		65-69	249	69655	357,48	4000		
19		70-74	200	43815	456,46	3000		
20		75-79	253	44536	568,08	2000		
21		80-84	159	23986	662,89	1000		
22		85+	68	12844	529,43	1000		
23		Total				100000		
24								

In Figure 7, we can verify if the result of calculation of the specific death rate for the age group 40-44 is correct.

2. Step 2 - calculation of the crude rate for total population.
   Prior calculating the crude rate for total population, the totals of number of cases and population need to be calculated. Figure 8 presents the procedure for calculating the totals by using the SUM function.

Figure 8. The procedure of calculation of totals for number of deaths and for Croatian population

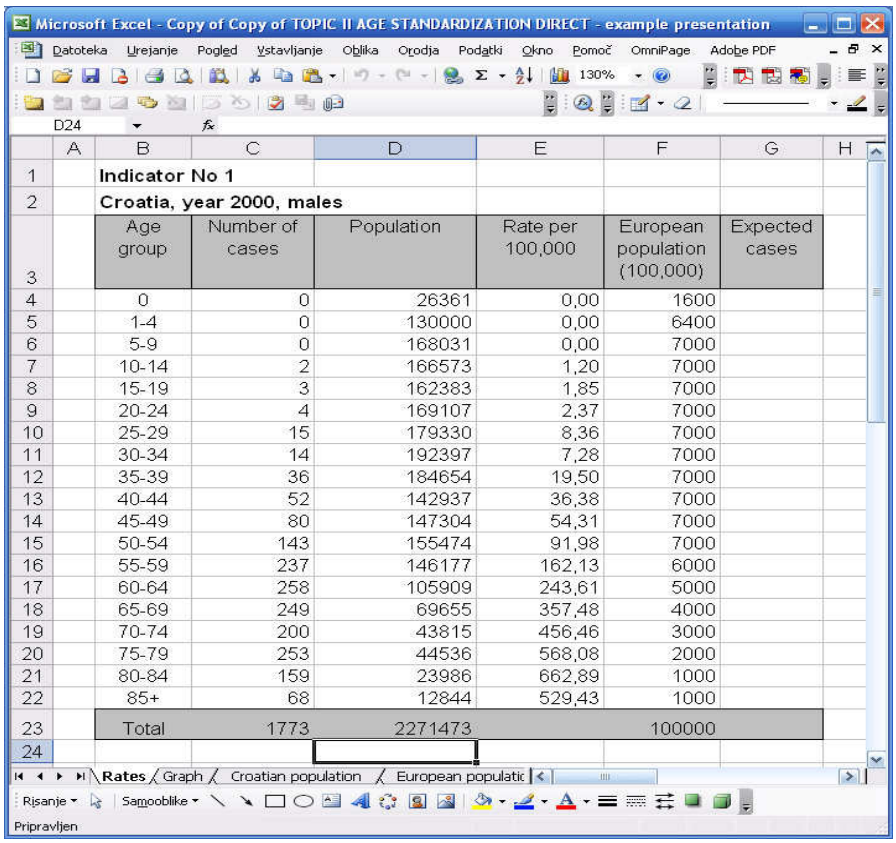


In Figure 9 the results of this procedure are presented. By comparing the totals in Tables 2 and 3 we can verify if they are correct. In continuation, the crude rate for total population is calculated by using the Equation 2 (Equation 6).

$$_{\text{crude}}\text{DR}_{(\text{total pop})} = \frac{1,773}{227,1473} \times 100,000 = 78.06 \qquad \textbf{Equation 6.}$$

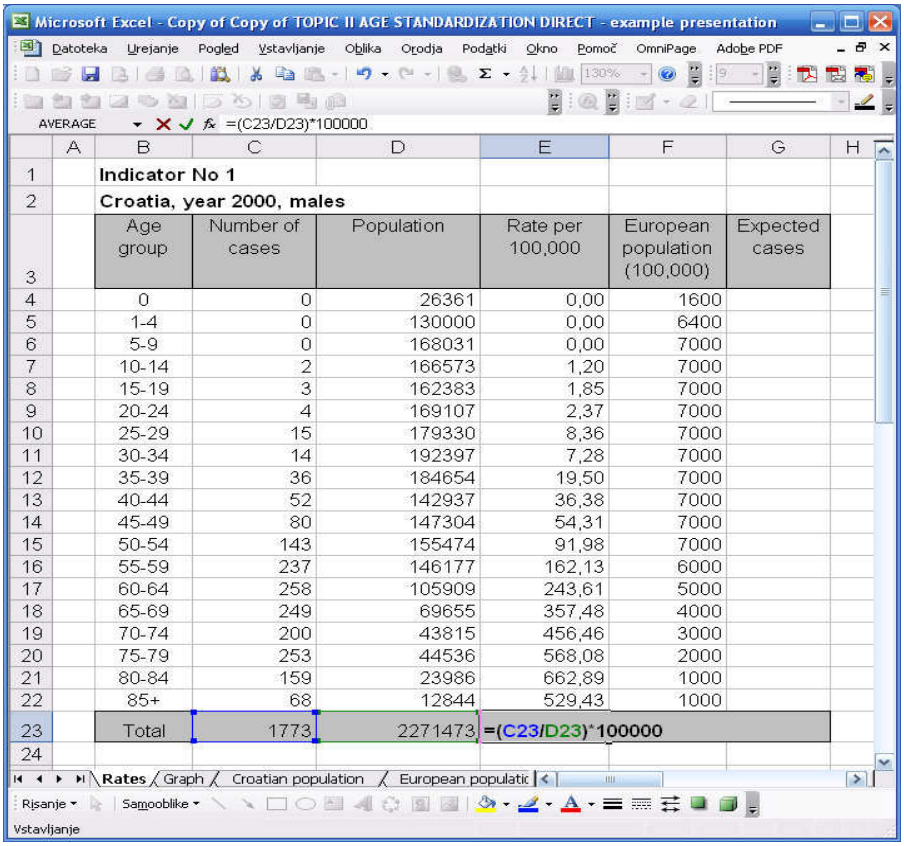
The procedure of calculation of the crude rate for total population using corresponding cells for number of deaths and the observed (actual local) total population in the spreadsheet is presented in Figure 10, while the result in Figure 11.

Figure 9. The result of the procedure of calculation of totals for Croatian population, number of cases and European standard population



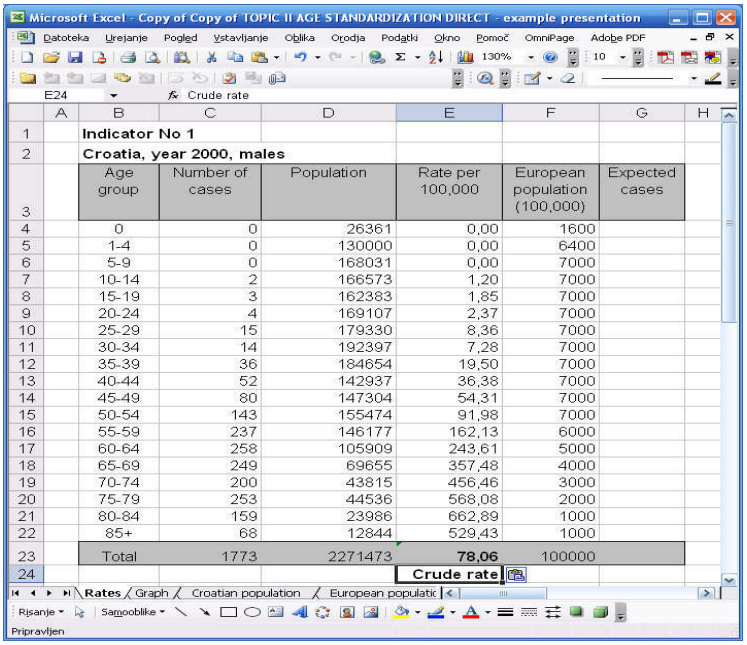
	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	0,00	1600		
5		1-4	0	130000	0,00	6400		
6		5-9	0	168031	0,00	7000		
7		10-14	2	166573	1,20	7000		
8		15-19	3	162383	1,85	7000		
9		20-24	4	169107	2,37	7000		
10		25-29	15	179330	8,36	7000		
11		30-34	14	192397	7,28	7000		
12		35-39	36	184654	19,50	7000		
13		40-44	52	142937	36,38	7000		
14		45-49	80	147304	54,31	7000		
15		50-54	143	155474	91,98	7000		
16		55-59	237	146177	162,13	6000		
17		60-64	258	105909	243,61	5000		
18		65-69	249	69655	357,48	4000		
19		70-74	200	43815	456,46	3000		
20		75-79	253	44536	568,08	2000		
21		80-84	159	23986	662,89	1000		
22		85+	68	12844	529,43	1000		
23		Total	1773	2271473		100000		
24								

Figure 10. The procedure of calculation of the crude rate for total population using corresponding cells for number of deaths and the observed (actual local) total population



	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	0,00	1600		
5		1-4	0	130000	0,00	6400		
6		5-9	0	168031	0,00	7000		
7		10-14	2	166573	1,20	7000		
8		15-19	3	162383	1,85	7000		
9		20-24	4	169107	2,37	7000		
10		25-29	15	179330	8,36	7000		
11		30-34	14	192397	7,28	7000		
12		35-39	36	184654	19,50	7000		
13		40-44	52	142937	36,38	7000		
14		45-49	80	147304	54,31	7000		
15		50-54	143	155474	91,98	7000		
16		55-59	237	146177	162,13	6000		
17		60-64	258	105909	243,61	5000		
18		65-69	249	69655	357,48	4000		
19		70-74	200	43815	456,46	3000		
20		75-79	253	44536	568,08	2000		
21		80-84	159	23986	662,89	1000		
22		85+	68	12844	529,43	1000		
23		Total	1773	2271473	=(C23/D23)*100000			
24								

Figure 11. The result of calculation of the crude rate for total population Calculation of the specific rate for age group 0



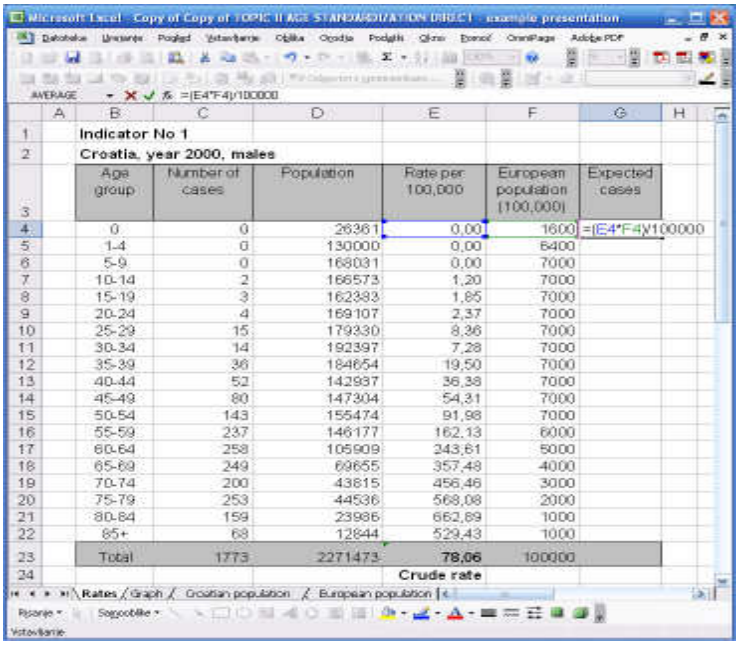
	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	0,00	1600		
5		1-4	0	130000	0,00	6400		
6		5-9	0	168031	0,00	7000		
7		10-14	2	166573	1,20	7000		
8		15-19	3	162383	1,85	7000		
9		20-24	4	169107	2,37	7000		
10		25-29	15	179330	8,36	7000		
11		30-34	14	192397	7,28	7000		
12		35-39	36	184654	19,50	7000		
13		40-44	52	142937	36,38	7000		
14		45-49	80	147304	54,31	7000		
15		50-54	143	155474	91,98	7000		
16		55-59	237	146177	162,13	6000		
17		60-64	258	105909	243,61	5000		
18		65-69	249	69655	357,48	4000		
19		70-74	200	43815	456,46	3000		
20		75-79	253	44536	568,08	2000		
21		80-84	159	23986	662,89	1000		
22		85+	68	12844	529,43	1000		
23		Total	1773	2271473	78,06	100000		
24					Crude rate			

3. Step 3 - calculation of the expected number of deaths in the standard population for every specific age group.
In the next step, the expected number of deaths in a specific age group is calculated by using Equation 3. In Figure 12 the equation for calculating the expected number of deaths in the standard population for the age group 0 using corresponding cells for crude death rate and the standard population in this age group in a spreadsheet is presented. In Figure 13 the results of this step in the procedure is presented.
In Equation 7 the procedure for calculating the expected number of deaths in the standard population for the age group 40-44 is presented, as well as the result.

$$N_{expected_{age40\text{-}44}} = \frac{36.38 \times 7{,}000}{100{,}000} = 2.55 \qquad \textbf{Equation 7.}$$

In continuation, the formula from the cell containing the function for calculating the crude specific death rate for the age group 0 is copied to other cells in the same column by dragging the right lower corner of the cell and extending it to the last age group cell (Figure 14).

Figure 12. Calculation of the number of expected number of deaths in the standard population for age group 0 by using corresponding cells for crude death rate and the standard population in this age group in a spreadsheet



	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	0,00	1600	=E4*F4/100000	
5		1-4	0	130000	0,00	6400		
6		5-9	0	168031	0,00	7000		
7		10-14	2	166573	1,20	7000		
8		15-19	3	162383	1,85	7000		
9		20-24	4	169107	2,37	7000		
10		25-29	15	179330	8,36	7000		
11		30-34	14	192397	7,28	7000		
12		35-39	36	184654	19,50	7000		
13		40-44	52	142937	36,38	7000		
14		45-49	80	147304	54,31	7000		
15		50-54	143	155474	91,98	7000		
16		55-59	237	146177	162,13	6000		
17		60-64	258	105909	243,61	5000		
18		65-69	249	69655	357,48	4000		
19		70-74	200	43815	456,46	3000		
20		75-79	253	44536	568,08	2000		
21		80-84	159	23986	662,89	1000		
22		85+	68	12844	529,43	1000		
23		Total	1773	2271473	78,06	100000		
24					Crude rate			

Figure 13. The result of calculation of expected number of deaths in the standard population for age group 0

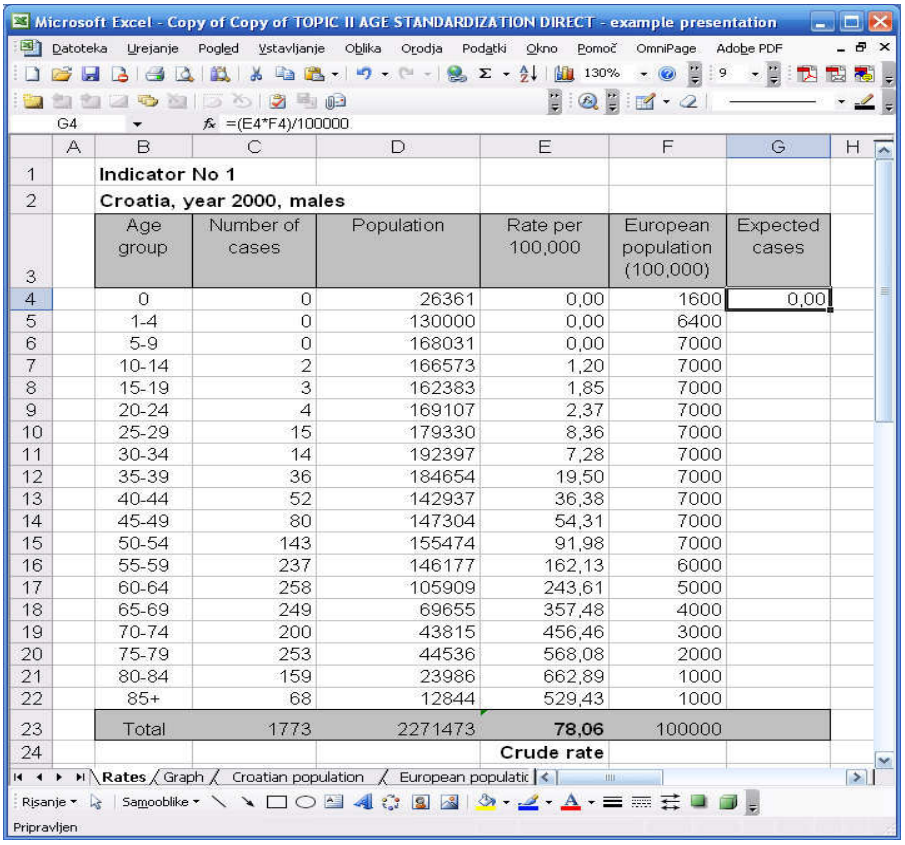


Figure 14. Copying of the function used for calculation of the number of expected number of deaths in the standard population for age group 0 to all age groups by dragging the right lower corner of the cell and extending it to the last age group cell

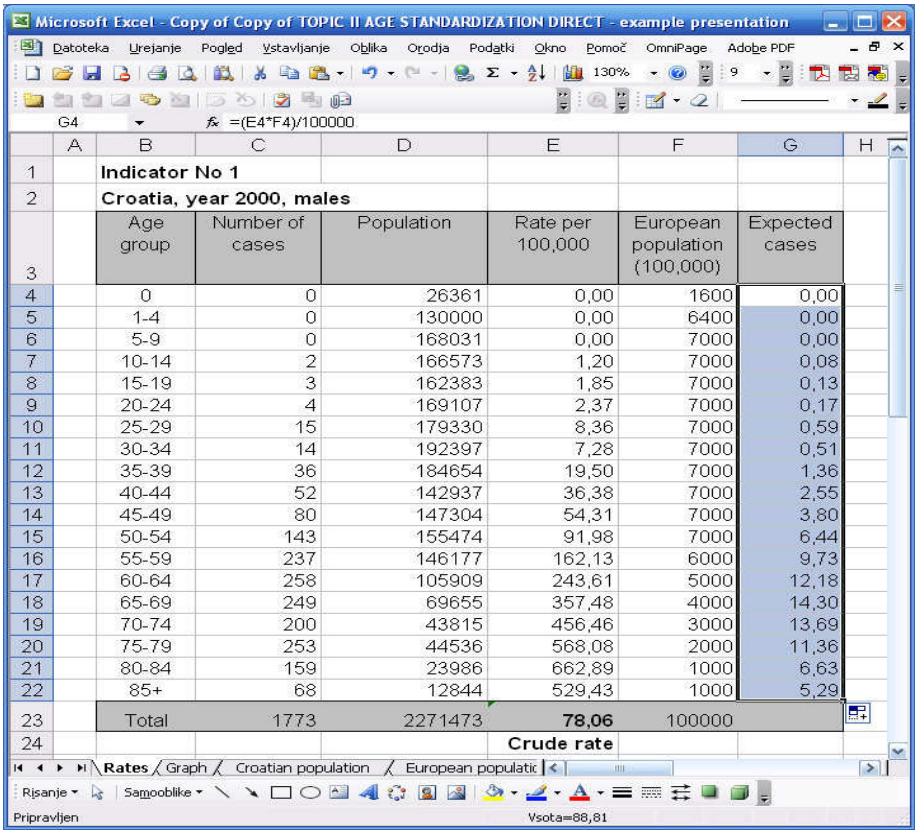
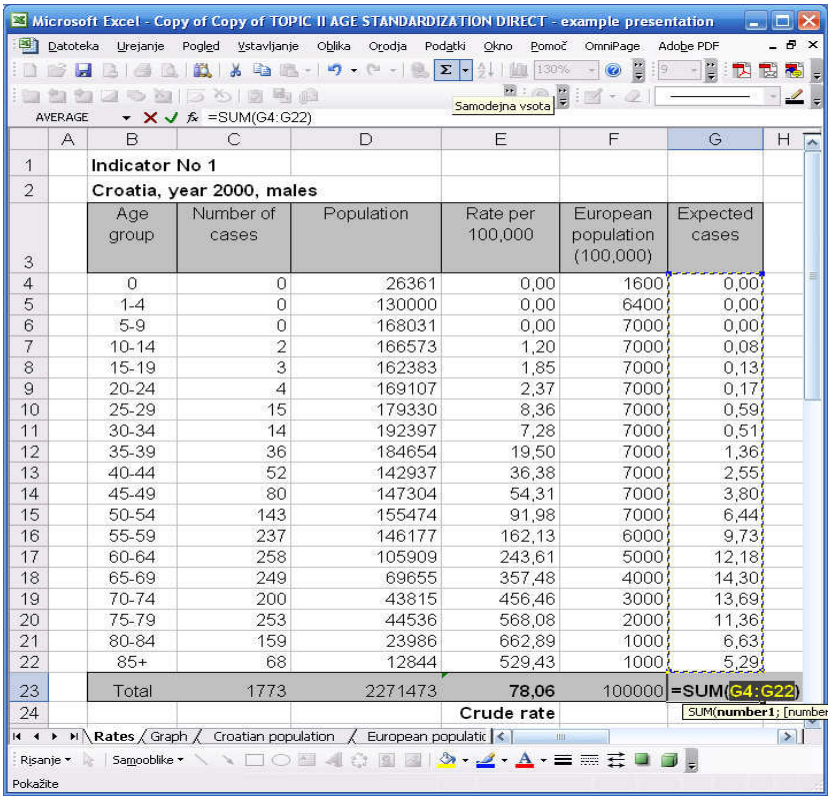
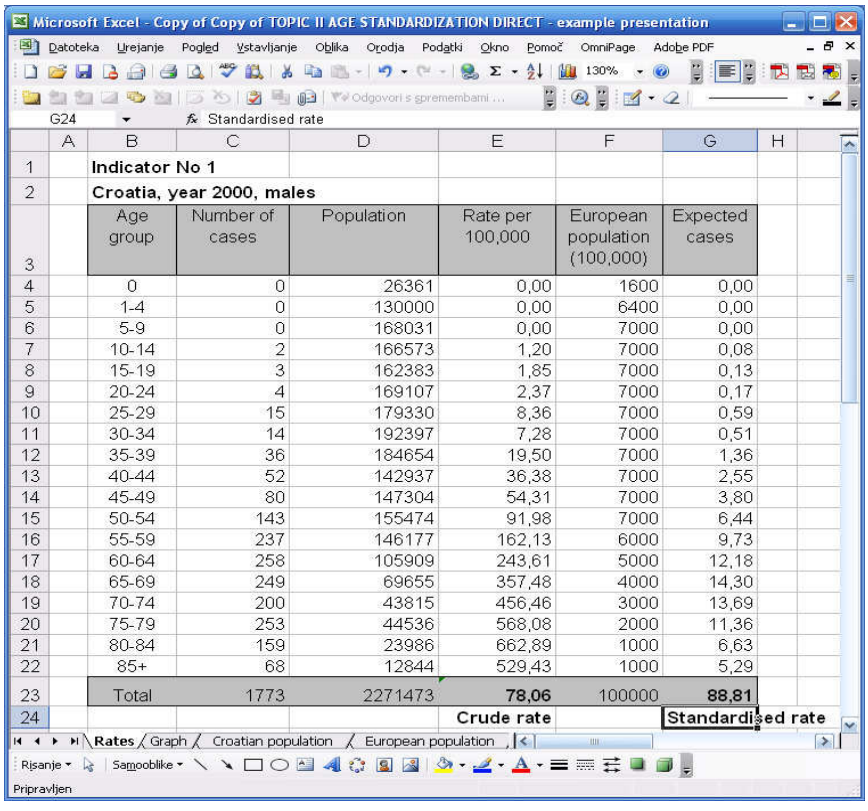


Figure 15. The procedure of calculation of the crude rate for total population using the SUM function in MS Excel programme



4. Step 4 - calculation of the standardized death rate in a total population.
   Finally, the standardized death rate is obtained by summation of number of expected cases in standard population across all age groups. Figure 15 presents the procedure for calculating the totals by using the SUM function in MS Excel programme, while Figure 16 presents the final result of the procedure.

Figure 16. Final result of the procedure of calculating standardized death rate



Exercises

Task 1

Read carefully the theoretical background of this module and discuss the confounding phenomenon with other students.

Task 2

Compare and interpret the crude rate and the SDR for disease D for male population of Croatia (Figure 16) given the disease is cancer (all sites). What do you think such a result mean?

Task 3

Perform the standardization procedure for female population for the disease D mortality in Croatia. Number of death cases (absolute incidence) is presented in Table 4. Follow the procedure presented in this paper from Step 1 to Step 4 (Figures 3 thru 16)[5].

Table 4. Number of death cases (absolute incidence) of the disease D in Croatia for female population for every age group for year 2000 (8). Source: National Health Institute of Croatia

Age group	Number of cases	Age group	Number of cases	Age group	Number of cases
0	0	30-34	10	65-69	270
1-4	0	35-39	29	70-74	240
5-9	0	40-44	24	75-79	330
10-14	1	45-49	67	80-84	238
15-19	0	50-54	136	85+	102
20-24	2	55-59	158	Total	1825
25-29	9	60-64	209		

Task 4

Compare:

- your results obtained in the Task 3 to results of other students,
- the results of female part of the population (Task 3) to the male part of the population (Case study),
- try critically to discuss the differences.

Task 5

Critically discuss strengths and limitations of standardization procedure in controlling the confounding phenomena.

References

1. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: May 31, 2013.
2. Centers for Disease Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Disease Control and Prevention; [cited 2009 Jun 25]. Available from URL: www.cdc.gov/osels/scientific_edu/ss1978/SS1978.pdf. Accessed: May 31, 2013.
3. Last JM, editor. A Dictionary of epidemiology. Fourth Edition. New York, New York: Oxford University Press; 2001.
4. Gordis L. Epidemiology. Second Edition. Philadelphia: W.B. Saunders Company; 2000.
5. Dos Santos Silva I. Dealing with confounding in the analysis. In: Dos Santos Silva I. Confounding Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. p.305-332. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: May 31, 2013.
6. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
7. Perace N, Greenland S. Confounding and interaction. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.371-97.

[5] The answer:
Crude death rate: 74.80
Standardized death rate: 62.53

8. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
9. Milic N. Standardization. In: Kirch W, editor. Encyclopedia of public health. Leipzig: Springer Verlag; 2008. p.1335-8.
10. World Health Organization, Regional Office for Europe. European health for all database. Copenhagen: World Health Organization, Regional Office for Europe, 2002. Available from URL: http://www.euro.who.int/en/what-we-do/data-and-evidence/databases/european-health-for-all-database-hfa-db2. Accessed: May 31, 2013.
11. Cancer Registry of the Republic of Slovenia [homepage on the Internet]. Available from URL: http://www.onko-i.si/rrs. Accessed: June 4, 2013.
12. Statistical Office of the Republic of Slovenia. Statistical Yearbooks of the Republic of Slovenia [homepage on the Internet]. Available from URL: http://www.stat.si/eng/pub_letopis_prva.asp. Accessed: June 4, 2013.
13. U.S. Census Bureau. International Data Base Entry [homepage on the Internet]. Available from URL: http://www.census.gov/population/international/data/idb/informationGateway.php. Accessed: June 4, 2013.
14. Kramer S. Clinical epidemiology and biostatistics. A primer for clinical investigators and decision-makers. Berlin: Springer-Verlag; 1988.
15. World Health Organization, Regional Office for Europe. European health for all database. Manual. Copenhagen: World Health Organization, Regional Office for Europe, 2002. Available from URL: ftp://ftp.euro.who.int/hfa/hfa-db.pdf. Accessed: September 5, 2009.
16. U.S. National Institutes of Health, National Cancer Institute, Surveillance, Epidemiology and End Results (SEER) Program. Standard Populations (Millions) for Age-Adjustment [homepage on the Internet]. Available from URL: http://seer.cancer.gov/stdpopulations/index.html. Accessed: June 4, 2013.
17. Croatian National Institute of Public Health. Croatian health service yearbook 2000. Zagreb: Croatian National Institute of Public Health; 2001.
18. Central Bureau of Statistics of Republic of Croatia. Population by sex and age, by settlements. Census of population, households and dwellings, 2001. Zagreb: Central Bureau of Statistics of Republic of Croatia, 2001. Available from URL: http://www.dzs.hr/default_e.htm. Accessed: June 4, 2013.

Recommended readings

1. Centers for Disease Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Disease Control and Prevention; [cited 2009 Jun 25]. Available from URL: www.cdc.gov/osels/scientific_edu/ss1978/SS1978.pdf. Accessed: May 31, 2013.
2. Dos Santos Silva I. Dealing with confounding in the analysis. In: Dos Santos Silva I. Confounding Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. p.305-332. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: May 31, 2013.
3. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: May 31, 2013.

HEALTH INVESTIGATION: ANALYSIS – PLANNING – EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Measures of central tendency and dispersion**
Module: 2.5	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Gena Grancharova, MD, PhD, Associate Professor, Dean Faculty of Public Health, Medical University of Pleven, Bulgaria Silviya Aleksandrova, MD, MB, Assistant professor Faculty of Public Health, Medical University of Pleven, Bulgaria
Address for correspondence	Gena Grancharova Medical University of Pleven 1, Kliment Ochridski Str. Pleven 5800, Bulgaria e-mails: dean-ph@mu-pleven.bg; gegran@optisprint.net
Keywords	Inter-quartile range, mean, median, mode, standard deviation, variance.
Learning objectives	After completing this module students and public health professionals should be able to: • explain why summary measures are needed in medicine and public health; • define the meaning and compute the mean, median and mode of a given set of data (grouped and ungrouped); • define the meaning and compute the range, inter-quartile range, variance, standard deviation and coefficient of variation; • describe the limitations, advantages and disadvantages of different measures of central tendency and dispersion; • select appropriate measures of central tendency and dispersion for a given data situation; • discuss the concept of normality of health data in terms of measures of central tendency and dispersion.
Abstract	Central tendency and dispersion measures (CTDM) are essential for summarizing any data set of individual scores. This process is based on two main characteristics of quantitave data - its variability and tendency to some typical levels. This section is devoted to the numerical approach of data summarising with the objective to underline the meaning of different measures, to present some simple basic methods of converting the raw data into meaningful summary statistics. The advantages and disadvantages of different CTDM are underlined in relation to different scales of data presentation, and to different forms of frequency distributions. The module is also describing the main uses of different CTDM and the concept of normality of health data.
Teaching methods	Two-hour lecture introduces the students to the main concepts of CTDM. After the lecture, students read and discuss in groups all the material presented in this section and individually answer the multiple choice questionnaire.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (1-2 students), LCD projector, access to the Internet and statistical package software; • training materials: recommended readings or other related readings; • target audience: bachelor and master students in public health.
Assessment of students	Multiple choice questionnaire (MCQ) - minimum 70% success.

MEASURES OF CENTRAL TENDENCY AND DISPERSION

Gena Grancharova, Silviya Aleksandrova

Theoretical background

Introduction

In the chapter on organizing and describing data it was demonstrated how raw data could be organized and presented in a meaningful way. A frequency distribution gives a good general picture of the pattern of the observations but sets of measurements cannot be adequately described only by the values of all individual measurements. For many purposes, the overall summary of a group's characteristic is of utmost importance (1,2). The researcher usually asks questions such as "What is the average fruit consumption in the group under study?", or "What is the average age of women delivering low birth weight infants?" Such questions seek a single number that best represents the whole distribution of the corresponding data values. It is obvious that further summarization is necessary, particularly before inferences or generalizations are drawn from the data under observation (3-6).

The process of summarization is based on two main characteristics of quantitave data.

1. firstly, the principal problem encountered in working with medical and public health data is its variability. Even when we follow all the strict requirements for random sampling we would almost certainly obtain different values of the variables studied in particular populations and samples, and
2. secondly, despite the individual fluctuations, many of the variables used in the behavioural or life sciences are distributed in such a way that most scores fall in the middle, with fewer scores falling on either side, in the "tails" of the distribution (7). In other words, the values of the most quantitative variables tend to some typical "middle" level (central point, or the most characteristic value) around which all the values are distributed. Because an index of typicalness is more likely to be representative if it comes from the center of a distribution than if it comes from either extreme, such measures are referred to as measures of central tendency (2,6,8).

The central tendency is due to determining factors and causes inherent in all cases of a given sample or population while the variability or dispersion is due to specific factors which may occur in some cases, but may be absent in others.

There are two basic methods of summarization: numerical and graphical. The objective of the numerical approach is to convert masses of numbers (raw data) into meaningful summary statistics (indices), reduced to a single number, that convey information about the average (typical) degree of a given variable and the degree to which observations differ (the degree of dispersion or spread).

Measures of central tendency

Before presenting the specific measures of central tendency, it is important to be familiar with some basic terms. An array of a set of numbers is simply those numbers in (algebraically) ordered sequence from the lowest to the highest (5). Each array has the following basic components:

- x - each individual raw score in a sample or in a population;
- n - the number of cases in a sample;
- N - the number of cases in a population;
- f - frequency (the number of observations with the same value);
- range - the difference between the largest and the smallest value in an array;
- Σx - the sum of all values in a sample or in a population

Before presenting the specific measures of central tendency, it is important also to know the shape of the distribution (there are distributions that do not assume a "normal" distribution) and the dispersion of the scores in order to interpret the data correctly.

Measures of central tendency are measures of the location of the middle or the center of a distribution. The definition of "middle" or "center" is purposely left somewhat vague so that the term "central tendency" can refer to a wide variety of measures. The three most commonly reported measures of central tendency are (9,10):

- the arithmetic mean,
- the median, and
- the mode.

Mean

Arithmetic mean

The mean is the most commonly used measure of central tendency. When the word "mean" is used without a modifier, it can be assumed that it refers to the arithmetic mean.

In a sample, arithmetic mean is denoted by $\bar{x}$, whereas in a population it is denoted by μ.

Mathematically, the mean is the sum of all the scores divided by the number of scores. In Equation 1 calculation of mean in a sample is presented as follows:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum x}{n} \qquad \textbf{Equation 1.}$$

$\bar{x}$ = mean of a sample
n = number of units in a sample studied
x_n = value of a single unit

If we are studying a population, then the calculation is exactly the same only denotation is different (Equation 2):

$$\mu = \frac{x_1 + x_2 + \ldots + x_n}{N} = \frac{\sum x}{N} \qquad \textbf{Equation 2.}$$

μ = mean of a population
N = number of units in a population studied
x_n = value of a single unit

As the public health investigations are mostly based on samples from which the conclusions and generalizations for the populations are made, further on in the text we will use $\bar{x}$ as a symbol of arithmetic mean.

The approaches to compute the arithmetic mean depend on the way on which the initial data are presented (raw or groupped data), and, if we have no statistical computer programme or hand calculator available, on the number of cases (statistical units) (3).

Calculation of the arithmetic mean in practice when raw data are available is presented in Case study 1, Case 1. The procedure of calculation has already been presented in Equations 1 and 2.

When only grouped data are available, the mean can also be calculated. The procedure is different when group intervals are equal, or when they are not:

- groups of unequal intervals (Equation 3):

$$\bar{x} = \frac{\sum xf}{\sum f} \qquad \textbf{Equation 3.}$$

$\bar{x}$ = arithmetic mean
$\sum xf$ = the sum of products of values of variable X (x_1, x_2, x_3, ... x_n) and the absolute frequency f for each value of variable X
$\sum f$ = the sum of absolute frequency f for each value of variable X that equals to the number of cases

Calculation of the arithmetic mean in practice when only grouped data are available, and intervals of groups are not equal is presented in Case study 1, Case 2.

This approach could be quite tedious when the sample is big and the variable has many values. In such situations, the transformation of the grouped data into equal width intervals (classes) is preferred;

- groups of equal intervals (classes) (Equation 4):

$$\bar{x} = \frac{\sum cf}{\sum f} \qquad \textbf{Equation 4.}$$

$\bar{x}$ = arithmetic mean
c = the mid-point of the class interval
$\sum f$ = the number of cases
$\sum cf$ = is the sum of the products of c and Σf

Calculation of the arithmetic mean in practice when only grouped data are available, and intervals of groups (classes) are equal is presented in Case study 1, Case 3.

Weighted arithmetic mean

When the research design includes observation of a sample consisting of two or more groups with different number of cases in each group, the weighted arithmetic mean would be the most appropriate measure of central tendency for the whole sample (6,11,12).

This approach is similar to the computing of the mean for grouped data but, instead of the interval mid-points, the true group means are used in the calculation. Thus, the weighted mean takes into account the different number of cases (that is, unequal weight) and the real mean for each group. The calculation of the weighted mean follows two steps:

- firstly, we have to multiply the mean for each group by the corresponding number of cases in each group, and add up the totals;
- secondly, the sum of totals (obtained in the first step) is divided by the total number of cases in the sample.

Data values with larger weights contribute more to the weighted mean and data values with smaller weights contribute less to the weighted mean. The formula for calculation is as follows (Equation 5):

$$\bar{x}_w = \frac{\sum wx}{\sum w} \qquad \textbf{Equation 5.}$$

$\bar{x}_w$ = weighted arithmetic mean

w = weight assigned to each data value

Calcualtion of the weighted arithmetic mean in practice is presented in Case study 1, Case 4.

Characteristics of the mean

Finally, we consider it is worthwhile to summarize and present basic characteristics of the aritmetic mean:

1. The arithmetic mean is unquestionably the most widely used measure of central tendency. It has the advantage to substitute by one single number all the individual values of a given variable and describe the typical level of a variable in a data set.
2. For normal or roughly symmetric distribution, the mean is the most efficient and therefore the least subject to sample fluctuations of all measures of central tendency. But the mean can be misleading in skewed distributions since it can be greatly influenced by scores in the tail. Therefore, other statistics such as the median may be more informative for skewed distributions. Also, the geometric mean is less affected by extreme values than is the arithmetic mean and is useful as a measure of central tendency for some positively skewed distributions.
3. Substantial disadvantage of the mean is that it can be drastically affected by the presence of a small number of outliers, e.g. observed values that are strikingly different from the rest (either unusually large or small). Such extreme values can distort the mean, particularly if there is a small number of subjects. Nonetheless, there are methods that allow to eliminate some extreme values and compute a new mean, which will be more typical for a particular empirical distribution. Such method is the criterion U, calculated as the ratio of the difference between the outlier (x_i) and the mean $\bar{x}$ and the standard deviation s. The computed criterion U is then compared with the table of critical values of u_t and if $u \geq u_t$, the extreme value x_i is disgarded as unusual.
4. The sum of the deviations of the scores in the distribution from the mean always is equal to zero. This is true because, by definition, the mean is the mathematical centre of the data. Thus, half of the distribution is above and half below the mean.
5. The sum of the squares of the deviations around the mean is smaller than the sum of squares around any other value. This characteristics of the mean underlies the calculation of the "least squares", which is used in applying some other statistical methods, such as regression analysis.
6. If to each value of the frequency distribution the same number is added or subtracted, then the mean increases or decreases by the same number.
7. The mean is not generally a "real" value and this makes the acceptance and interpretation of the data sometimes more difficult – for example, a mean number of children in a sample might be 2.4, or an average number of limbs 3.997.

Median

The median (Me) is the measure of central tendency which is not calculated, but identified or determined. It is a member of a family of measures of location called quantiles that are described in detail in a separate module. It is a measure that informs about the value of the middle observation when data have been arranged in ordered series from the lowest to the highest value. Half of units of a

sample (or population) lie above the median and half below the median. The procedure for identifying the median is:

- rearrange all observations (units) in order of magnitude (from the smallest to the largest) in an ordered series (one must be sure to list all data values even though some values may repeat more than once);
- then we must determine whether the number of cases is odd or even;
- when the number of observations is odd, the median is simply the value of the middle observation (unit) in the ordered series;
- when the number of observations is even, the median is just a halfway of values of the two middle observations.

Identification of the median in practice is presented in Case study 1, Case 5.

Because the median is less sensitive than the mean to the biasing influence of extreme scores on a data set, it is best used when a distribution is known to be asymmetric or when its shape is otherwise unknown.

The median is particalarly suitable for scales of measurement having ordinal characteristics and when the validity of assumptions about the size of the intervals between data points is questionable (9,13).

Characteristics of the median

Characteristics of the median are as follows:

1. The median is usually a realistic value, or at worst measured in half-units (when the number of observations is even).
2. A better advantage over the mean is that the median is more robust towards outliers (extreme scores). The presence of a few extreme observations in the sample (from whatever cause) do not affect the middle values (8). This makes the median a better measure than the mean for highly skewed distributions.
3. The only disadvantage of the median is that it does not include all the individual values of a variable. It reflects only one value in odd number of cases or two values in even number of cases.
4. The median is preferred measure of central tendency when (11):
    - the lowest and highest values of a quantitative variable are too far from the rest of the values;
    - there is uncertainty in some values;
    - it is not possible to determine the exact shape of the distribution, or when the distribution is highly asymmetric;
    - the number of cases is small.

Mode

The mode (Mo) is the observation in an array with the highest frequency of occurrence. The advantage of the mode as a measure of central tendency is that its meaning is obvious and it is the simplest to determine of the three measures of central tendency. Actually, the mode is not computed, but rather it is determined through inspection of a frequency distribution.

Although it is common for most distributions to contain exactly one mode (as in a normal distribution and large homogenious samples), it is possible for more than one mode to exist. A distribution having one mode is called unimodal. A distribution having two modes is called bimodal.

Identification of the mode in practice is presented in Case study 1, Case 6.

Characteristics of the mode

Characteristics of the mode are as follows:

1. The mode is a quick and easy method of determining the most popular score at a glance.
2. It is the only measure of central tendency that can be used with nominal data.
3. The mode is the weakest measure of central tendency. It often provides a crude and limited representation of the characteristics of a distribution as compared to the mean and the median. This is true because, in some cases, the mode may be the lowest or the highest value in the distribution.
4. It is greatly subject to sample fluctuations and is rather unstable, e.g. the modes tend to fluctuate widely from one sample drawn from a population to another sample drawn from the same population. Therefore, it is not recommended to be used as the only measure of central tendency. The mode is seldom used in research reports except in association with other measures of central tendency.

5. As the weakest measure of central tendency, the use of the mode is restricted to nominal scales of measurement and is seldom reported except in association with other measures of central tendency.
6. A further disadvantage of the mode is that many distributions have more than one mode. These distributions are called multimodal.
7. Nevertheless, the mode has a true meaning and this is very important in medicine and public health. For example, it is more important to determine which group has higher risk for some disease, e.g. to determine the mode in the age distribution instead of calculating the mean age of persons with the disease.

Other measures of central tendency

Trimmed mean

A trimmed mean is a systematic method for avoiding outliers when calculating means by discarding or "trimming off" a certain percentage of the lowest and the highest scores and then computing the mean of the remaining scores (14,15). After removing the specified observations, the trimmed mean is found using an arithmetic averaging formula. For example, a mean trimmed 50% is computed by discarding the lower and higher 25% of the scores and taking the mean of the remaining scores.

A trimmed mean is obviously less susceptible to the effects of extreme scores than is the arithmetic mean. It is therefore less susceptible to sampling fluctuations than the mean for extremely skewed distributions. It is less efficient than the mean for normal distributions. This method is best suited for data with large, erratic deviations, or extremely skewed distributions.

A trimmed mean is stated as a mean trimmed by X%, where X is the sum of the percentage of observations removed from both the upper and the lower bounds. For example, to trim the mean by 40% it means that we remove the lowest 20% and the highest 20% of the values.

In contrast to the arithmetic mean, the trimmed mean is a robust measure of central tendency. For example, a small fraction of anomalous measurements with abnormally large deviation from the center may change the mean value substantially. At the same time, the trimmed mean is stable in respect to the presence of such abnormal extreme values, which get "trimmed" away.

Trimmed means are often used in Olympic scoring to minimize the effects of extreme ratings possibly caused by biased judges, where the extreme scores are often discarded before computing the score for a particular performance.

Calculation of trimmed mean in practice is presented in Case study 1, Example 7.

Trimean

The trimean is a measure of central tendency computed by using quantiles (percentiles or quartiles and median) that are described in detail in a separate module in this book. If we use percentiles for calculation, the procedure is as follows (Equation 6):

- adding the 25^{th} percentile,
- plus twice the 50^{th} percentile,
- plus the 75^{th} percentile, and
- dividing by four.

$$TM = \frac{P_{25} + (2 \times P_{50}) + P_{75}}{4} \qquad \textbf{Equation 6.}$$

TM = trimean
P_{25} = 25^{th} percentile
P_{50} = 50^{th} percentile
P_{75} = 75^{th} percentile

Exactly the same results we obtain if we use the quartiles and median) for calculation (Equation 7):

$$TM = \frac{Q_1 + (2 \times Q_2) + Q_3}{4} = \frac{Q_1 + (2 \times Me) + Q_3}{4} \qquad \textbf{Equation 7.}$$

TM = trimean
Q_1 = 1^{st} quartile
Q_2 = 2^{nd} quartile
Q_3 = 3^{rd} quartile
Me = median

The trimean is a good measure of central tendency but it is not used as much as it should be (16). It is almost as resistant to extreme scores as the median and is less subject to sampling fluctuations

than the arithmetic mean in extremely skewed distributions. It is less efficient than the mean for normal distributions.

Calculation of trimean in practice is presented in Case study 1, Example 8.

Geometric mean

The geometric mean is a type of mean or average, which indicates the central tendency or typical value of a set of numbers. It is similar to the arithmetic mean, except that instead of adding the set of numbers and then dividing the sum by the count of numbers in the set, n, the numbers are multiplied and then the n^{th} root of the resulting product is taken. It is rarely used in biomedical research. However, it should be obligatory applied when the values of the variable increase in geometric progression, or the distribution of frequencies by the logarithms of observed values of the variable are approximately symmetric (17).

The geometric mean only applies to positive numbers. It is also often used for a set of numbers whose values are meant to be multiplied together or are exponential in nature, such as data on the growth of the human population.

Calculation of the geometric mean is easy – it is just the n^{th} root of the product of the scores (Equation 8) (16):

$$GM = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots \times x_n} = \left(x_1 \times x_2 \times x_3 \times \ldots \times x_n\right)^{\frac{1}{n}} = \left(\prod x\right)^{\frac{1}{n}} \qquad \textbf{Equation 8.}$$

GM = geometric mean
n = number of units in a sample studied
x_n = value of a single unit
Πx = product of all scores

The geometric mean can also be computed by taking the logarithm of each number, or computing the arithmetic mean of the logarithms.

Calculation of the geometric mean in practice is presented in Case study 1, Example 9.

The geometric mean must be used when working with percentages (which are derived from values), whereas the standard arithmetic mean works with the values themselves (16).

Comparison of the measures of central tendency

Of the five measures of central tendency, the mean is the most stable. If repeated samples were drawn from a given population, the means would vary or fluctuate less than the modes or medians. Because of its stability, the mean is the most reliable estimate of the central tendency of the population.

The arithmetic mean is the most appropriate measure in situations in which the concern is for totals of combined performance of a group. When the primary concern is to learn what a typical value is, then the median would be preferred (2).

Of the five measures of central tendency discussed, the mean is by far the most widely used. It takes every score into account, is the most efficient measure of central tendency for normal distributions and is mathematically tractable making it possible for statisticians to develop statistical procedures for drawing inferences about means. On the other hand, the mean is not appropriate for highly skewed distributions and is less efficient than other measures of central tendency when extreme scores are possible. The geometric mean is a viable alternative if all the scores are positive and the distribution has a positive skew.

The median is useful because its meaning is clear and it is more efficient than the mean in highly-skewed distributions. However, it ignores many scores and is generally less efficient than the mean, the trimean, and trimmed means.

The mode can be informative but should almost never be used as the only measure of central tendency since it is highly susceptible to sampling fluctuations.

The trimean and trimmed means are both examples of statistics developed to resist sampling fluctuations. It is highly recommended that at least one of these two be computed in addition to the mean (14,15,18).

The level of measurement plays an important role in determining the appropriate index of central tendency that can be used to describe a variable. In general, the mode is appropriate for nominal measures. The median is appropriate for ordinal measures. The mean is appropriate for interval and ratio measures (7).

However, the higher the level of measurement, the greater the flexibility we have in choosing a descriptive statistic. Variables measured on an interval or ratio scale can use any of the three measures of central tendency, although it is usually preferable to use the mean (2).

In skewed distributions, the values of the mode, median and mean differ. The mean is always pulled in the direction of the long tail (17). This means that the mean is typically higher than the median and the mode in positively skewed distributions (Mo<Me< $\bar{x}$) (Figure 1). The mean is lower than the median and the mode in negatively skewed distributions (Mo>Me> $\bar{x}$) (Figure 2) (3,17). When a distribution is

symmetric and unimodal, the three measures of central tendency – the mean, the median and the mode – coincide (3,17).

Figure 1. Arithmetic mean ($\bar{x}$), mode (Mo) and median (Me) in a positively skewed distribution

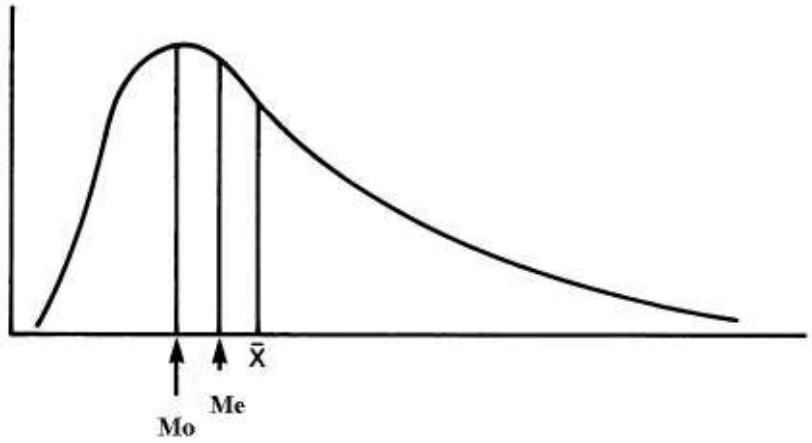
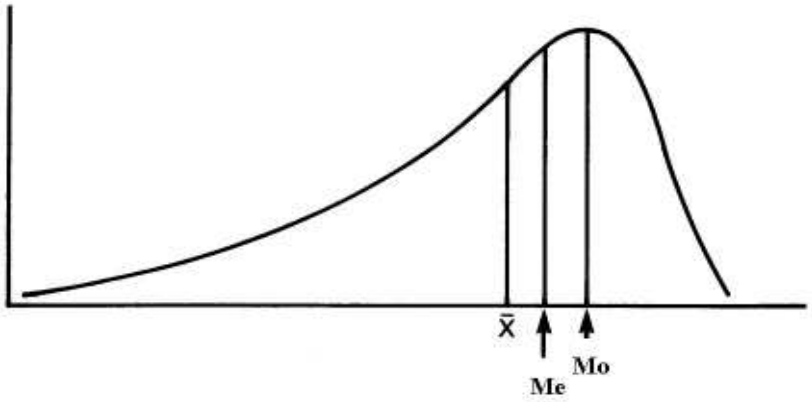


Figure 2. Arithmetic mean ($\bar{x}$), mode (Mo) and median (Me) in a negatively skewed distribution



To illustrate the difference between arithmetic and geometric mean, the best is to think in terms what question answer both of them. While arithmetic mean answers the question, "If all the units in a sample (population) had the same value, what would that value have to be in order to achieve the same total?", the geometric mean answers the question, "If all the units in a sample (population) had the same value, what would that value have to be in order to achieve the same product?"

Measures of dispersion (variability)

Why do we need measures of dispersion?

We need measures of dispersion because measures of central tendency do not give a total picture of a distribution.

Two sets of data with identical means could be different from one another. Two distributions with the same means could be very different in shape: for example, they could be skewed in opposite directions (2). Secondly, even when two sets of data have equal means, medians, modes, and the same form of distribution, they could be different from one another. Knowledge of a single summary figure for describing the location of a center of a sample or population is not enough without a measure of the extent of variability or spread of the measurements around this summary index. Illustration is given in Case study 2, Case 1.

Health workers often have to decide whether to classify an individual as healthy or sick, suffering from a particular disease or not, needing treatment or not, etc. For this task, the co-called "normal" values of certain measurements provide the necessary yardstick. But, the word "normal" value is a statistical concept and depends, to a great extent, on the distribution of the classifying attribute in the population. Measures of spread or dispersion or variability are, therefore, a complete description of a given health data set. No description of any health data by summary measures is

complete without the measures of variability (6). Several such measures have been developed, the most common of which are:

- the range,
- standard deviation,
- variance,
- inter- and semiquartile range, and
- coefficient of variation.

Range

The range is simply the difference between the extreme values (the highest and the lowest) of the variable in a given empiric distribution (12,19). It is usually denoted by d (difference) and is expressed as follows (Equation 9):

$$d = x_{max} - x_{min} \qquad \textbf{Equation 9.}$$

d = the range

x_{max} = the lowest value

x_{max} = the highest value

The main virtue of the range is the ease with which it can be computed.
As an index of variability, the shortcomings of the range outweigh this modest advantage. The range being based on only two scores (and the two most unusual ones), is a highly unstable index. The drawback to this simple measure is also the fact that a single outlier may have a large impact on the range (8) (Case study 2, Case 2).

Another difficulty with the range is that it ignores completely variations in scores between the two extremes. Surely, a value that seeks to measure variation within a group of individuals should reasonably be expected to be based on information gathered from all the individuals under study, not merely on two selected and unrepresentative ones.

For these reasons, the range is used only as a gross descriptive index and is typically reported in conjunction with, not instead of, other measures of variability (6,10,13,20).

Standard deviation and variance

The standard deviation (denoted by SD or s for a sample and σ for a population) is the most commonly reported measure of variability, especially with interval- or ratio-level data. It is a statistic that describes the degree of variation among the individual observartions in the sample (21). Like the mean, the standard deviation considers every score in a distribution. For this reason, means and standard deviations are generally reported together, whether in the text or in tables (7).

The standard deviation represents the average deviation of scores in a given distribution around the central score, e.g. the mean, which serves as a reference point or baseline for the entire data set. The calculation of SD includes the following steps (3,10):

- firstly, we need to calculate how much each individual varies from the mean by simply subtracting the mean from each individual value: $\mathbf{x}-\bar{x}$ **;**
- measuring the overall variation present in the study group by adding the individual variations together - this sum is denoted by $\Sigma(x-\bar{x})$. In order to calculate the average deviation, the sum $\Sigma(x-\bar{x})$ should be divided by the number of scores **n.** We would expect this to be large if the variation is large, and small if most individuals are very similar to the mean value and hence show relatively little variation. Unfortunately, totalling the differences of deviations in this way tells us nothing at all. The reason is that those individuals who have values larger than the mean (e.g. $\mathbf{x}-\bar{x}$ is a positive value) will simply cancel out those who have values below the mean (e.g. $\mathbf{x}-\bar{x}$ is a negative value). So, any data set, highly consistent or highly variable, will result in a zero value when all the deviations are added together: $\Sigma(x-\bar{x})$ will always be equal to zero;
- to get around the problem of positive and negative variation, we can square each difference so that both positive and negative values end up as positive and no longer cancel out one another. The new measure of the variation in a group of individuals is therefore $\Sigma(x-\bar{x})^2$ and it is referred to as the sum of squares, or the sum of squared deviations from the mean;
- the sum of squares as a measure of variation has some limitations related to the number of results under study. To allow fair comparisons between studies of different sizes, and to provide a truly universal measure of variation, it is reasonable to take the study size into account by calculating an "average" variation. This measure is called variance (s^2) and is calculated as follows: (Equation 10):

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$ **Equation 10.**

s^2 = variance of a sample
x = value of a single unit
$\bar{x}$ = arithmetic mean of a sample
n = number of units in a sample studied
n-1 = degrees of freedom (df)

- finally, the variance s^2 has one major drawback as a truly useful measure of variation. It is based on squared differences, and hence it measures variation in squared units which is not convenient (for example, it is nonsense to say that the variation in blood pressure study is in squared millimeters of mercury). To solve this problem we need to take the square root of the variance and we will come to the most meaningful and most widely used measure of variability known as the standard deviation - s or SD (Equation 11):

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$ **Equation 11.**

s = standard deviation of a sample
x = value of a single unit
$\bar{x}$ = arithmetic mean of a sample
n = number of units in a sample studied
n-1 = degrees of freedom (df)

Calcualtion of variance and standard deviation in practice is presented in Case study 2, Case 3.

The standard deviation can also be used in interpreting individual scores from within a distribution. Using the basic principle of normal distribution (Figure 3), we can determine the limits of different groups of normality to assess and interpret the individual results.

When the distribution of scores is normal or nearly normal, it is possible to say even more about the standard deviation. There are about 3 SDs above and 3 SDs below the mean with normally distributed data. Furthermore, in a normal distribution, as shown in Figure 3, a fixed percentage of cases falls within certain distances from the mean (2,3):

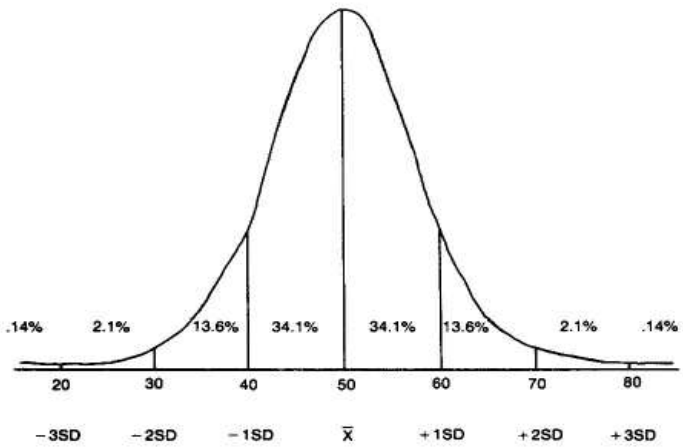
- 68% percent of all cases fall within 1 SD of the mean (34% above and 34% below the mean);
- 95% of the scores fall within 2 SDs from the mean;
- only a handful of cases – about 2% at each extreme – lay more than 2 SDs from the mean.

Using this principle, we can easily create "normal limits" and interpret individual scores for many clinical and laboratory tests.

In case of skewed distributions, the "normal limits" can be established based on percentiles, as described in a separate module in this book.

In summary, the standard deviation is a useful index of variability that can be used to describe an important characteristic of a distribution and also can be used to interpret the score of the performance of an individual in relation to others in the sample. Like the mean, the standard deviation is a stable estimate of a population parameter and also is used in more advanced statistical procedures. The standard deviation is the preferred measure of distribution's variability but it is appropriate only for variables measures on the interval or ratio scale (2,3,22,23).

Figure 3. Standard deviations in a normal distribution



Interquartile and semiquartile ranges

The interquartile range

The interquartile range (IQR) is the difference between the third (Q_3) and the first (Q_1) quartiles in a dataset (where quartiles are the values that divide the data into four equal sized parts). Quartiles are described in detail in a separate module. Characteristics of IQR are:

- the advantage of the IQR over the range is that it is quite robust to outliers;
- the IQR is commonly quoted in conjunction with the sample median.

Calcualtion of IQR in practice is presented in Case study 2, Case 4.

The semiquartile range

The semiquartile range (SQR), used as a term in many statistical texts instead of IQR, is half of the distance between Q_1 and Q_3.

Because these two measures of variability are based on middle cases rather than extreme scores, they are considerably more stable than the range (2,8).

Coefficient of variation

The standard deviation s and the variance s^2 have the same measurement units as the mean $\bar{x}$ and because of this they are not appropriate for comparing the relative variability of different distributions where the variables are measured in different units (height, weight, blood pressure, cholesterol level, etc.) This problem can be overcome by calculating another measure of variation called the coefficient of variation (denoted by CV), also known as relative variability. It expresses the sample standard deviation in a proportion or percentage of the mean value (1). When expressed as percentage it is calculated by the formula (Equation 12):

$$CV = \frac{s}{\bar{x}} \times 100 \qquad \textbf{Equation 12.}$$

CV = coefficient of variation
$\bar{x}$ = arithmetic mean of a sample
s = standard deviation of a sample

The main advantage of the coefficient of variation is its independence of any unit of measurement, and thus useful for comparison of variability in two or more distributions having variables expressed in different units (1,6). For example, if we measure height and weight in a sample, it is not possible to say which variable varies greatly because these two variables have different measurement units. Using the coefficient of variation we can transform the standard variations in comparable units, expressed in percents. The interpretation is as follows (3):

- when the value of the coefficient of variation is less than 10% we could say that the degree of variation is low and the sample is quite homogeneous;
- in a situation when 10%<V<30% - the variation is moderate;
- with V>30% the variation is considerable, and this is a clear evidence of heterogeneity of the sample or population under study.

Calcualtion of coefficient of variation in practice is presented in Case study 2, Case 4.

Summary

In this module, two essential measures of location for describing and representing frequency distributions were discussed: measures of central tendency and variability.

The measures of central tendency outlined were the mode and the median for discrete data, and the mean for continuous data. The weighted mean, trimean and trimmed mean were shortly presented as well.

Measures of variability were shown to be the range, the variance, the standard deviation, the inter- and semiquartile range, and the coefficient of variation.

All these statistics are appropriate for processing the data to the point that a distribution of raw data can be meaningfully represented by only two statistics. We have seen that the mean and the standard deviation are the most appropriate for interval or ratio data when the distribution is similar enough to normal distribution. The median and the inter- or semiquartile range are used when the data is measured on an ordinal scale, or when interval or ratio data is found to have a highly skewed distribution.

The content of the module focused on the meaning and the use of these concepts, rather than stressing on calculations that nowadays are made usually by any statistical package software.

Case study

Case study 1: Measures of central tendency

Case 1 – calculation of arithmetic mean when raw data are available

Suppose the age at first birth for a sample of 10 mothers is (24):

18 21 23 23 25 27 27 28 30 33

Then, the mean age is given according to Equation 1 by (Equation 13):

$$\bar{x} = \frac{18 + 21 + 23 + \ldots + 33}{10} = \frac{255}{10} = 25.5 \qquad \textbf{Equation 13.}$$

Case 2 – calculation of arithmetic mean when only grouped data are available and intervals of groups are of unequal size

Suppose the values of height for 100 female urban liveborns are those presented in Table 1.

Table 1. Elements of calculation of arithmetic mean when only groupped data are available and intervals of groups are of unequal size

Height in cm *(x)*	Frequency (f)	Product *(x*f)
46	2	92
47	6	282
48	7	336
49	20	980
50	30	1,500
51	20	1,020
52	8	416
53	5	265
54	2	108
	Σf = n = 100	Σ*x*f = 4,999

The arithmetic mean can be computed according to Equation 3 through the following algorithm (Equation 14):

- each value of the variable is multiplied by its frequency and the product xf is recorded in the appropriate row and column;
- the products in column 3 are summed and recorded as Σ*x*f;
- the sum Σ*x*f is divided by the number of cases (Σf = n) to come to the mean $\bar{x}$.

$$\bar{x} = \frac{4{,}999}{100} = 49.99 = 50 \qquad \textbf{Equation 14.}$$

Case 3 – calculation of arithmetic mean when only grouped data are available and intervals of groups are of equal size

Let us take the same example of height for 100 female urban liveborns and regroup the data into three equal class intervals (Table 2).

Table 2. Elements of calculation of arithmetic mean when only groupped data are available and intervals of groups are of equal size

Height in cm *(x)*	Frequency (f)	Mid-point of the class interval (c)	Product (cf)
46 - 48	15	47	705
49 - 51	70	50	3,500
52 - 54	15	53	795
	Σf = n = 100		Σcf = 5,000

The algorithm for computing the mean is, according to Equation 4, as follows (Equation 15):

- defining the interval width;
- regrouping the data into equal width intervals and summing the frequences for each interval;
- defining the mid-point of each interval;
- multiplying the frequency of each interval by its mid-point;
- recording the products cf in the corresponding rows;
- summing the products cf and recording the sum Σcf in the bottom line;
- dividing the sum Σcf by the number of cases Σf to come to $\bar{x}$.

$$\bar{x} = \frac{5{,}000}{100} = 50.00 = 50 \qquad \textbf{Equation 15.}$$

It is obvious that the arithmetic mean is the same as in case 1, but with intervals the computing is easier.

Let us now consider another the set of data presented in Table 3 - Systolic blood pressure for 240 men (data modified from Lwanga et al. (6).

Table 3. Elements of calculation of arithmetic mean when only grouped data are available and intervals of groups are of equal size

Systolic blood pressure in mmHg (class interval) *(x)*	Frequency (f)	Mid-value of the class (c)	Product (cf)
90-99,9	4	95	380
100-109,9	16	105	1,680
110-119,9	18	115	2,070
120-129,9	40	125	5,000
130-139,9	66	135	8,910
140-149,9	56	145	8,120
150-159,9	34	155	5,270
160- 169,9	6	165	990
Total	$\Sigma f = n = 240$	-	$\Sigma cf = 32,420$

The arithmetic mean is then calculated according to Equation 4 as follows (Equation 16):

$$\bar{x} = \frac{32,420}{240} = 135.1 \qquad \textbf{Equation 16.}$$

Case 4 - calcualtion of the weighted arithmetic mean

Suppose mean ages of preschool children in 5 different villiages (data modifeid from Lwanga et al. (6)) (Table 4). On the basis of available data we can, according to Equation 5, calculate weighted arithmetic mean (Equation 17).

Table 4. Elements of calculation of weighted arithmetic mean

Village No.	No. of children (w)	Mean age (in months) (x)	Product (wx)
1	35	61.5	2,152.5
2	40	62.0	2,480.0
3	45	62.5	2,812.5
4	50	63.5	3,175.0
	$\Sigma w = n = 170$		$\Sigma wx = 10,620$

$$\bar{x}_w = \frac{10,620}{170} = 62.47 \qquad \textbf{Equation 17.}$$

Thus, the weighted mean is 62.47 months.

Case 5 - identification of the median

Suppose the age at first birth for a sample of 10 mothers is (8):

18 21 23 23 25 27 27 28 30 33

In this set of data where the number of units in the sample is even (ten), the median age at first birth is 26 – the mean of the two middle numbers 25 and 27.

18 21 23 23 25 27 27 28 30 33

In continuation, we add to the observed set of data one more case:

18 21 23 23 25 27 27 28 30 33 33

In this set of data where the number of units in the sample is now odd (eleven), the median age at first birth is 27 - simply the value of the middle observation in the ordered series of data:

18 21 23 23 25 27 27 28 30 33 33

Case 6 - identification of the mode

Suppose the age at first birth for a sample of 10 mothers is (8):

18 21 23 23 25 27 27 28 30 33

In this set of data, the variable “age at first birth” has two modes - 23 and 27.

18 21 23 23 25 27 27 28 30 33

In continuation, we change the set of data: the case number 3 has value 22 instead of 23:

18 21 22 23 25 27 27 28 30 33

In this set of data, the variable "age at first birth" has one mode - 27.

18 21 22 23 25 27 27 28 30 33

Case 7 - calculation of the trimmed mean

As a first example, we have calculation of mean score of referees in a figure skating competition. A figure skating competition produces the following scores:

6.0 8.1 8.3 9.1 9.9

A mean trimmed 40% by using Equation 1 and trimming 40% of values would equal (Equation 18):

$$\bar{x}_{trimmed} = \frac{8.1 + 8.3 + 9.1}{3} = 8.5 \qquad \textbf{Equation 18.}$$

whereas the ordinary mean would be (Equation 19):

$$\bar{x} = \frac{6.0 + 8.1 + 8.3 + 9.1 + 9.9}{3} = 8.28 \qquad \textbf{Equation 19.}$$

We could see that the trimmed mean is larger than the ordinary arithmetic mean. To trim the mean by 40%, we remove the lowest 20% and highest 20% of values, eliminating the scores of 6.0 and 9.1. As shown by this example, trimming the mean can reduce the effects of outlier bias in a sample (18).

Let us take now a set of final exam scores for a 40-question test of 20 students (modified from Weiss (14,15). The results of the test are as follows:

2 15 16 16 19 21 21 25 26 27 4 15 16 17 20 21 24 25 27 28

The following steps in calculating the trimmed mean include:

1. Exploration of the data in order to find the outliers. For accomplishing this step we need first to arrange the data into an ordered series:

   2 4 15 15 16 16 16 17 19 20 21 21 21 24 25 25 26 27 27 28

   In this example there are two outliers - the low values of 2 and 4.
2. Computation of the usual mean for the data (Equation 20):

$$\bar{x} = \frac{2 + 4 + 15 + \ldots + 27 + 27 + 28}{20} = 19.3 \qquad \textbf{Equation 20.}$$

3. Computation of the 10% trimmed mean for the data (Equation 21):

$$\bar{x}_{trimmed} = \frac{15 + 15 + 16 + \ldots + 27 + 27 + 28}{20} = 20.2 \qquad \textbf{Equation 21.}$$

4. If we compare the two means obtained we can conclude that the trimmed mean provides a better measure of central tendency for this set of data.

Case 8 - identification of the trimean

Let's take the values for the following 50 measurements (23):

50 50 50 50 50 50 50 50 50 50
51 51 51 51 51 51 51 51 52 53
53 53 55 55 55, 55 56 56 56 58
58 59 60 60 61 63 63 63 64 67
67 69 70 70 75 77 78 80 85 103

The P_{25} (Q1), P_{50} (Q2), and P_{75} (Q3) of this dataset are 51, 55, and 63 respectively. Therefore, the trimean is computed according to Equation 6 as follows (Equation 22):

$$TM = \frac{51 + (2 \times 55) + 63}{4} = 56 \qquad \textbf{Equation 22.}$$

Case 9 - calculation of the geometric mean

The geometric mean of two numbers, say 2 and 8, would be according to Equation 8 just the square root (i.e., the second root) of their product (Equation 23):

$$GM = \sqrt{2 \times 8} = \sqrt{16} = 4 \qquad \textbf{Equation 23.}$$

As another example, the geometric mean of 1, ½, and ¼ is the cube root (i.e., the third root) of their product (0.125), which is ½ (16) (Equation 24).

$$GM = \sqrt[3]{1 \times 0.5 \times 0.25} = \sqrt[3]{0.125} = 0.5$$ **Equation 24.**

The geometric mean of the scores: 1, 2, 3, and 10 is the fourth root of their product (Equation 25):

$$GM = \sqrt[4]{1 \times 2 \times 3 \times 10} = \sqrt[4]{60} = 2.78$$ **Equation 25.**

Case study 2: Measures of dispersion

Case 1 – variability of the data

Consider two sets of data of mothers' ages at first birth. The first sample of 10 cases with age values is:

18, 21, 23, 23, 25, 27, 27, 28, 30, 33.

The second sample of 10 cases with age values is:

23, 23, 24, 25, 26, 26, 27, 27, 27, 27.

The means for the two samples are 25.5, the medians 26, and the modes 27.

However, it is obvious that the samples are quite different. The second sample seems more homogeneous that the first one.

It is unquestionable, from this simple example, that knowledge of a single summary figure for describing the characteristics of a sample or population is not enough without a measure of the extent of variability or spread of the measurements around this summary index.

Case 2 – the range

Consider two sets of data from Case study 2, Case 1, again:

Set No. 1: 18 21 23 23 25 27 27 28 30 33

Set No. 2: 23 23 24 25 26 26 27 27 27 27

The means for the two samples are 25.5, the medians 26, and the modes 27, while the range is very different. The range for the first sample equals (according to Equation 9) to 15, and for the second sample to 4. So, from sample to sample, drawn from the same population, the range tends to fluctuate considerably.

Case 3 – variance and standard deviation

Consider two samples from Case study 2, Case 1, again. In Table 5, elements for calculation of standard deviation are presented.

Table 5. Elements of calculation of standard deviation in two samples with equal arithmetic means

Sample 1				Sample 2			
Age (years) *(x)*	**Mean** ($\bar{x}$)	**Diff.** (x- $\bar{x}$)	**Square of diff.** (x- $\bar{x}$)^2	**Age (years)** *(x)*	**Mean** ($\bar{x}$)	**Diff.** (x- $\bar{x}$)	**Square of diff.** (x- $\bar{x}$)^2
18	25.5	-7,50	56,25	23	25.5	-2,50	6,25
21	25.5	-4,50	20,25	23	25.5	-2,50	6,25
23	25.5	-2,50	6,25	24	25.5	-1,50	2,25
23	25.5	-2,50	6,25	25	25.5	-0,50	0,25
25	25.5	-0,50	0,25	26	25.5	0,50	0,25
27	25.5	1,50	2,25	26	25.5	0,50	0,25
27	25.5	1,50	2,25	27	25.5	1,50	2,25
28	25.5	2,50	6,25	27	25.5	1,50	2,25
30	25.5	4,50	20,25	27	25.5	1,50	2,25
33	25.5	7,50	56,25	27	25.5	1,50	2,25
			Σ(x- $\bar{x}$)^2 = 176,5				Σ(x- $\bar{x}$)^2 = 24,5

Variance and standard deviation for the first sample are, according to Equations 10 and 11, as follows (Equations 26 and 27):

$$s^2_{sample1} = \frac{176.5}{9} = 19.6 \Rightarrow s_{sample1} = \sqrt{19.6} = 4.43$$ **Equation 26.**

Variance and standard deviation for the second sample are:

$$s^2_{sample2} = \frac{24.5}{9} = 2.7 \Rightarrow s_{sample2} = \sqrt{2.7} = 1.65$$ **Equation 27.**

These calculations confirm that the variation is much smaller in the second sample which means that the mean for the second sample is a more accurate measure of central tendency.

Case 4 – calculation of interquartile range

Let's go back to the example with two samples of 10 mothers at first birth:

Set No. 1: 18 21 23 23 25 27 27 28 30 33

For the first dataset $Q_1 = 23$ and $Q_3 = 28$, so the IQR is 5 years.

Set No. 2: 23 23 24 25 26 26 27 27 27 27

For the second data set $Q_1 = 24$ and $Q_3 = 27$, so the IQR is 3 years.

Case 5 – calculation of coefficient of variation

Let's go back to the example with two samples of 10 mothers at first birth again:

Set No. 1: 18 21 23 23 25 27 27 28 30 33

Set No. 2: 23 23 24 25 26 26 27 27 27 27

We have already calculated the mean and the standard deviation for both sets of data (Case study 2, case 3). The mean was in both cases the same (25.5 years), while the standard deviations were different ($s_1 = 4.43$ years; $s_2 = 1.65$ years). According to Equation 12, coefficient of variation could be calculated as follows (Equations 28 and 29):

$$CV_1 = \frac{4.45}{25.5} \times 100 = 17.45\% \qquad \textbf{Equation 28.}$$

$$CV_2 = \frac{1.65}{25.5} \times 100 = 6.47\% \qquad \textbf{Equation 29.}$$

We can conclude that the first sample has much higher variability than the second one. The second sample could be described as homogenous, while in the first sample the variation is moderate.

Exercises

Task 1

For the following questions choose between "true" or "false":

1. With nominal data, the mean should be used as a measure of central tendency.	T	F
2. The mode represents the most frequently occurring score in a distribution.	T	F
3. With ordinal data we can use both the mode and the mean as a measure of central tendency.	T	F
4. When the data are interval or ratio, we can use the mean as a measure of central tendency.	T	F
5. If a continuous distribution is highly skewed, the median might be the appropriate measure of central tendency.	T	F
6. When a frequency distribution is positively skewed, the mean is greater than the median or the mode.	T	F
7. Given a normal distribution, the three measures of central tendency are equivalent.	T	F
8. The range is the simplest indicator of variability.	T	F
9. The range is calculated by adding the lowest score to the highest score in a distribution.	T	F
10. The square root of the variance is called the standard deviation.	T	F
11. Standard deviation indicates the extent to which scores are distributed around the mean.	T	F
12. When a distribution consists of very different scores, standard deviation will be relatively large.	T	F
13 The median is less affected than the mean by extreme scores of a distribution.	T	F
14. Central tendency describes the 'typical' value of a set of scores.	T	F
15. If the number of raw scores is odd, the median is the score in the	T	F

middle position.

16. The mean must have a value equal to one of the scores in the distribution. T F

17. About 10% of scores fall 3 standard deviations above the mean. T F

Answers: 1. F; 2. T; 3. F; 4. T; 5. T; 6. T; 7. T; 8. T; 9. F; 10. T; 11. T; 12. T; 13 T; 14 T; 15 T; 16 F; 17F

Task 2

For the following two questions choose only one right answer:

1. Given a set of nominally scaled scores, the most appropriate measure of central tendency is the:
   A. Mean
   B. Mode
   C. Standard deviation
   D. Range
2. Which of the following statements is true?
   A. The mode is the most useful measure of central tendency.
   B. The variance is the square root of the standard deviation.
   C. The median and the 50^{th} percentile rank have different values.
   D. The mean is more affected by extreme scores than the median.

Answers: 1. B; 2. D

Task 3

The following four questions refer to the following set of data:

2 2 3 4 6 6 7.

1. Σ x is equal to:
   A. 30
   B. 40
   C. 50
   D. None of the above A, B or C.
2. $(\Sigma x)^2$ is equal to:
   A. 124
   B. 128
   C. 130
   D. 900
3. The median is equal to:
   A. 6
   B. 5
   C. 4
   D. 3
4. The range for the above set of scores is:
   A. 7
   B. 5
   C. 2
   D. 1

Answers: 1. A; 2. D; 3. C; 4. B

Task 4

The following three questions refer to the example below:

- a clinic had 50 patients attending in a month. The number of times each patient visited the clinic is given below in the form of a frequency distribution (Table 6):

Table 6. Elements of calculation of location measures in a distribution of patients attending a clinic

No of visits (x)	No of patients (f)
7	3
6	6
5	6
4	10
3	21
2	0

1. The total number of patients' visits was:
    A. 194
    B. 28
    C. 50
    D. None of the above
2. The mean number of visits per patient was:
    A. 3.89
    B. 3.50
    C. 1.00
    D. 3.84
3. The median number of visits was:
    A. 3.88
    B. 3.50
    C. 3.00
    D. 4.00

Answers: 27. A; 28. D; 29. D

Task 5

Answer the following questions:

1. The more dispersed, or spread out, a set of scores is:
    A. The greater the difference between the mean and the median
    B. The greater the value of the mode
    C. The greater the standard deviation
    D. The smaller the interquartile range
2. The mean height of a student group is 167 cm. Assuming that height is normally distributed this enables us to deduce that:
    A. Approximately half of all students are taller than 167 cm
    B. Being a student stunts your growth
    C. Approximately half of all students are shorter than 167 cm
    D. A and C
    E. None of the above
3. If we subtract the value of the mean from every score in a set of scores the sum of the remaining values will be:
    A. Impossible to determine
    B. Equal to the mean
    C. A measure of the dispersion around the mean
    D. Zero
    E. None of the above
4. Given a normally distributed continuous variable the best measure of central tendency is the:
    A. Mode
    B. Median
    C. Mean
    D. Standard deviation
    E. None of the above
5. If a distribution is negatively skewed, then:
    A. The median is greater than the mean
    B. The mode is greater than the median
    C. The mean is greater than the median
    D. Both A and B are true
    E. None of the above is true
6. In a normal distribution, the mean, the median and the mode:
    A. Always have the same value
    B. The mean has the higher value
    C. The mean has the lower value
    D. Have no particular relationship
    E. Cannot take the same value
7. Given the group of scores 1, 4, 4, 4, 7, it can be said of the mean, the median, and the mode that:
    A. The mean is larger than either the median or the mode
    B. All are the same

C. The median is larger than either the mean or the mode
D. All are different
E. The mode is larger than either the median or the mode

Answers: 1. C; 2. D; 3. D; 4. C; 5. D; 6. A; 7. B

Task 6

The following four questions refer to the example below:

A nurse recorded the number of analgesic preparations taken by patients in a surgical ward as follows:

5 2 8 2 3 2 4 12

1. The mode for this distribution is:
   A. 2
   B. 3
   C. 8
   D. There is no mode
2. The median is:
   A. 2.00
   B. 3.50
   C. 3.00
   D. 3.25
3. The mean is:
   A. 3.52
   B. 5.43
   C. 4.75
   D. 4.15
4. The range is:
   A. 9
   B. 10
   C. 12
   D. 2

Answers: 1. A; 2. B; 3. C; 4. B

Task 7

The following two questions refer to the data set below:

3 3 4 5 6 7 8 9 9 10 38 60.

1. The interquartile range is:
   A. 5.0
   B. 4.5
   C. 6.0
   D. 9.0
2. The percentage of cases falling between SD = -1 and SD = +1 is:
   A. 16.8%
   B. 33.6%
   C. 34.1%
   D. 68.3%

Answers: 1. A; 2. D

References

1. Cramer MS. Clinical epidemiology and biostatistics. Springer-Verlag Berlin Heidelberg, 1988.
2. Polit DF, Hungler BP. Nursing research. Principles and methods. 6th rev. ed. Lippincot Williams & Wilkins; 1999.
3. Grancharova G, Hristova P. Medical statistics (in Bulgarian). Pleven: Publishing House of the Medical University of Pleven; 2004.
4. Kirkwood BR, Sterne JAC. Essentials of medical statistics. 2nd rev. ed. Oxford: Blackwell Publishing Company, 2003.
5. Leaverton PE. A review of biostatistics. A program for self-instruction. 2nd rev. ed. Boston: Little Brown and Company; 1981.

6. Lwanga SK, Tye CY, Ayeni O. Teaching health statistics. Lessons and seminar outlines. 2nd rev. ed. Geneva: World Health Organization; 1999.
7. Munro BH, Visintainer MA, Page EB. Statistical methods for health care research. Philadelphia: J.B. Lippincott Company; 1986.
8. Palmer C. Lectures in biostatistics. Teaching Course in Modern Epidemiology. Varna: Medical University of Varna; 1994.
9. Campbell MJ, Machin D, Walters SJ. Medical statistics: A textbook for the health sciences. 4th rev. ed. Chichester West Sussex: John Wiley & Sons LTD; 2007.
10. Hassard TH. Understanding biostatistics. St. Louis, Missouri: Mosby Year Book; 1991.
11. Kalinov K. Statistical methods in behavioural and social sciences (in Bulgarian). Sofia: NBU Publishing House; 2001.
12. Wonnacott TH, Wonnacott RJ. Introductory statistics. 3rd rev. ed. New York: John Wiley & Sons; 1977.
13. Maxwell DL, Satake E. Research and statistical methods in communication disorders. Baltimore, Maryland: Williams & Wilkins; 1997.
14. Weiss NA. Elementary statistics. Instructor's Manual. Reading, Massachusetts: Addison-Wesley Publishing Company; 1989.
15. Lane DM. HyperStat Online Statistics Textbook. Trimmed mean. Available from URL: http://davidmlane.com/hyperstat/example_1.html. Accessed: April 20, 2009.
16. Forbes Digital Company. Investopedia. Geometric mean. Available from URL: http://www.investopedia.com/terms/g/geometricmean.asp. Accessed: April 20, 2009.
17. Ranchov G. Biostatistics and biomathematics - concepts, methods, applications (in Bulgarian). Sofia: Eco Print; 2008.
18. Statistics.com. Trimean. Available from URL: http://www.statistics.com/resources/glossary/t/trimmean.php. Accessed: April 20, 2009.
19. Morton RF, Hebel JR. A study guide to epidemiology and biostatistics (including 100 multiple-choice questions. Baltimore: University Park Press; 1983.
20. Polgar S, Thomas SA. Introduction to research in the health sciences. 2nd rev.ed. Singapore: Churchill Livingstone; 1991.
21. Bailar JC, Mosteller F. Medical uses of statistics. Waltham, Massachusetts: NEJM Books; 1986.
22. Abramson JH. Survey methods in community medicine. 4th rev. ed. Edinburgh London Melbourne and New York: Churchill Livingstone; 1990.
23. Ingelfinger JA, Mosteller F, Thibodeau LA, Ware JH. Biostatistics in clinical medicine. New York: Macmillan Publishing CO, Inc.; 1987.
24. Weiss NA. Elementary statistics. Reading, MA: Addison-Wesley Publishing Company; 1989.

Recommended readings

1. Campbell MJ, Machin D, Walters SJ. Medical statistics: A textbook for the health sciences. 4th rev. ed. Chichester West Sussex: John Wiley & Sons LTD; 2007.
2. Kirkwood BR, Sterne JAC. Essentials of Medical statistics. 2nd rev. ed. Oxford: Blackwell Publishing Company; 2003.
3. Lwanga SK, Tye CY, Ayeni O. Teaching health statistics. Lessons and seminar outlines. 2nd rev. ed. Geneva: World Health Organization; 1999.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Measures of location: quantiles**
Module: 2.6	ECTS (suggested): 0.1
Author(s), degrees, institution(s)	Gena Grancharova, MD, PhD, Associate Professor, Dean Faculty of Public Health, Medical University of Pleven, Bulgaria Silviya Aleksandrova, MD, MB, Assistant professor Faculty of Public Health, Medical University of Pleven, Bulgaria
Address for correspondence	Gena Grancharova Medical University of Pleven 1, Kliment Ochridski Str. Pleven 5800, Bulgaria e-mails: dean-ph@mu-pleven.bg; gegran@optisprint.net
Keywords	Median, percentiles, quantiles, quartiles.
Learning objectives	After completing this module students and public health professionals should be able to: • understand the concept of quantiles; • define and compute different types of quantiles; • explain the use of quartiles and percentiles to summarize health data;
Abstract	Measures of central tendency and dispersion are essential for summarizing any data set of individual scores. This process is based on two main characteristics of quantitave data – its variability and its tendency to some typical level. This section is devoted to quantiles as measures of location in the numerical approach of data summarising with the objective to present identification as well as methods of calculation of different types of quantiles, especially percentiles and quartiles.
Teaching methods	A half an hour lecture in the form of Power Point presentation should introduce the students to the main concepts of quantiles, types of quantiles and their use. After the lecture, students should read and discuss in groups all the material presented in this section and individually answer the multiple choice questionnaire at the end of this lecture.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (1-2 student), LCD projector, access to the Internet and statistical package software; • training materials: recommended readings or other related readings; • target audience: bachelor and master students in public health.
Assessment of students	Multiple choice questionnaire (MCQ) - minimum 70% success.

MEASURES OF LOCATION: QUANTILES

Gena Grancharova, Silviya Aleksandrova

Theoretical background

Quantiles

Quantiles are special measures of location that could be described as the points that divide the ordered series of data (from the lowest to the highest value) into subgroups of equal size regarding the number of units in a subgroup. In other words, quantiles are the data values marking the boundaries between consecutive subgroups ordered series of data (an array).

Types of quantiles

There exist several types of quantiles, dividing an array in different number of subgroups: e.g., terciles divide the distribution into three equal subgroups (called thirds), quartiles - into four subgroups (quarters), quintiles - into five (fifths), deciles - into ten (tenths), and centiles - into hundred (hundredths) (1).

The most frequently used quantiles are presented in Table 1, including their number and number of subgroups they are dividing the distribution.

Table 1. Most frequently used quantiles

Quantile/s	Number of equal parts of a distribution	Number of quantiles of this type
Median	2	1
Quartiles	4	3
Deciles	10	9
Percentiles	100	99

As shown in Table 1, in all types of presented quantiles their number is one less then the number of corresponding equal parts of a distribution they are dividing.

Estimation of quantiles

Quantiles are measures of location that are usually not calculated, but identified or determined. The procedure of identifying quantiles is as follows:

- first we need to rearrange all observations according to the magnitude of a value of a variable we are observing in an ordered series of data from the lowest to the highest value (NOTE: one must consider all observations even though some values may repeat more than once);
- then we must determine whether the number of cases is odd or even;
- when the number of observations is odd, the location of the very first quantile dividing the distribution in two parts equal in number of observations (the median) is simple - the median is simply the value of the middle observation in the ordered series of data. When the number of observations is even, we locate the central two observations. Afterwards, we sum the values of these two units and divide the sum by two - the median is just the halfway of values of the two middle observations;
- then we repeat exactly the same procedure in the lower half and in the upper half of the ordered series to locate the quantiles dividing the ordered series in four equal parts regarding the number of units or observations (quartiles);
- we repeat the procedure until we divide the distribution in the required number of equal parts.

For the estimation of quantiles there exist also procedures to calculate them. They are described later in this module.

Median

Median (Me) is a quantile, dividing an ordered series into two equal halfs regarding the number of observations (Table 1). Usually it is identified from the ordered series using the following procedure:

- first we need to rearrange all observations according to the magnitude of a value of a variable we are observing in an ordered series of data from the lowest to the highest value (NOTE: one must consider all observations even though some values may repeat more than once);
- then we must determine whether the number of cases is odd or even;
- when the number of observations is odd, the median is simply the value of the middle observation (unit) in the ordered series of data;

- when the number of observations is even, the median is just a halfway of values of the two middle observations.

This simple procedure of dentification of the median in practice is presented in Case study, Example 1.

To identify the median in a long ordered series of data we can use the formula for calculating the position (rank) of the unit carrying median value (Equation 1).

$$Me_{Rank} = \frac{(n+1)}{2} \qquad \textbf{Equation 1.}$$

Me = rank of the unit carrying median value
n = number of cases

This procedure is also presented in Case study, Example 1.

Calculation of median

Median is usually identified but it could also be calculated if we have groupped data. For a grouped frequency distribution, the calculation of the median might be a little more complicated. If we assume that the variable is continuous (for example time, height, weight, or level of pain), we can use a formula for calculating the mean. This formula can be applied to ordinal data, provided the variable being measured has an underlying continuity (Equation 2) (2).

$$Me = X_L + \left(i \times \frac{(n/2) - cumf_L}{f_i} \right) \qquad \textbf{Equation 2.}$$

Me = median
X_L = number of observations below a given observation
i = width of the class interval
n = number of cases
$cumf_L$ = cumulative frequency at the real lower limit of the interval
f_i = frequency of cases in the interval containing the median

Calculation of the median in practice is presented in Case study, Example 2.

Use of median

It is used as a measure of central tendency. It is described in detail in separate module in this book.

Percentiles

Percentiles [also called centiles (1)] are points that divide an array into 100 equal parts (3-6). Thus, there are 99 percentiles. They are denoted as P_1, P_2, ... P_{25},...P_{50},...., P_{75}, ...P_{99}.

Some characterristics of the percentiles are as follows:

- a percentile tells us the relative position of a given observation,
- it allows us to compare scores on tests that have different means and standard deviations (7,8),
- the tenth percentile, for example, exceeds 10% and is exceeded by 90% of the observations. The seventy-fifth percentile exceeds 75% of the data, etc.

Calculation of percentiles

Percentiles are usually identified/determined, but they could also be calculated.

In an ungrouped data set, a percentile (P_i) can be calculated by the following formula (Equation 3):

$$P_i = \frac{n_{belowobsx}}{n_{all}} \times 100 \qquad \textbf{Equation 3.}$$

P_i = percentile i
$n_{below\ obs\ x}$ = number of observations below a given observation
n_{all} = number of all observations

Calculation of a percentile in practice in this case is presented in Case study, Example 3.

Calculation of percentiles in an interval scale is much more complicated. The following formula can be used (Equation 4) (3):

$$P_i = L_{pi} + \frac{c}{f} \times e \qquad \textbf{Equation 4.}$$

P_i = the value of the percentile i
L_{pi} = the lower limit of the interval where the percentile is situated
c = the difference between the percentile rank and the cumulative frequency in the previous interval
f = the number of cases in the percentile interval

e = the width of the percentile interval

An intermediate step in this procedure is the calculation of the rank of a given percentile (Equation 5):

$$\text{Percentile rank} = \frac{\sum f}{100} \times i \qquad \textbf{Equation 5.}$$

Σf = the total number of cases
i = the number, corresponding to the percentile rank (for P_{10} i = 10, for P_{25} - 25, etc.).

Calculation of a percentile in practice in this case is presented in Case study, Example 4.

Use of percentiles

Percentiles are practically used to establish the reference limits of normality in many clinical and other areas of investigation. For this purpose, usually seven main percentiles are used - P_3, P_{10}, P_{25}, P_{50}, P_{75}, P_{90} and P_{97} – to form the upper and lower limits of seven reference groups of a population.

Percentiles are widely used for the establishment of "normal ranges" of values for health data that permits the selection of appropriate actions in medical practice, or to allow for accurate estimates of many clinical and laboratory indicators. Percentiles have an advantage compared to the other methods of determining "normal" values as they are applicable to any form of variables/distribution (not only to normal distribution). When the investigator prefers to use seven reference groups the limits of "normal" values are determined by P_{25} and P_{75}, whereas P_{50} corresponds to the mean.

Quartiles

These are observations in an array that divide the distribution into four equal parts. Therefore, there 3 quartiles, denoted as Q_1, Q_2 and Q_3. If we have an array of 23 cases, the first quartile Q_1 is the 6^{th} observation with its corresponding value; the second quartile Q_2 is the 12^{th} observation, and the third quartile Q_3 is equal to the 18^{th} observation.

Identification of the quartiles in practice is presented in Case study, Example 5.

Comparison between different types of quantiles

It is worth mentioning that:

- P_{25} corresponds to Q_1,
- P_{50} corersponds to Q_2 and to the median, and
- P_{75} is equal to Q_3,
- Taking into account that the sample median (Me) is the second quartile (Q_2), then the median of the lower half of the data gives the first quartile (Q_1), and similarly, the median of the upper half of the data gives the third quartile (Q_3).

Use of quantiles

Quantiles are used in description of both, central tendency and dispersion of a distribution they are describing. Median is used as a measure of central tendency, while quartiles are used for a quick estimation of the degree of dispersion in an array.

Case study

Example 1 - identification of the median in an ungroupped data set

Suppose the age at first birth for a sample of 10 mothers is (9):

18 21 23 23 25 27 27 28 30 33

In this set of data where the number of units in the sample is even (ten), the median age at first birth is 26 – the mean of the two middle numbers 25 and 27.

18 21 23 23 25 27 27 28 30 33

In continuation we add to the observed set of data one case more:

18 21 23 23 25 27 27 28 30 33 33

In this set of data where the number of units in the sample is now odd (eleven), the median age at first birth is 27 – simply the value of the middle observation in the ordered series of data:

18 21 23 23 25 27 27 28 30 33 33

Now let us take two very simple examples to present the procedure of identification of the median by calculating the rank of the unit carrying median value first (Equation 1). Suppose we have the values of a variable in a group of 7 units (n = 7) are as follows:

5 8 9 10 15 18 28

Using the Equation 1 we find that in this ordered array the median will be the 4th score (Equation 6):

$$Me_{Rank} = \frac{(7+1)}{2} = 4 \qquad \textbf{Equation 6.}$$

So, the median is equal to 10. Now we have odd number of units. Suppose we have the the values of a variable a group of 8 units (n = 8) are as follows:

6 12 17 19 20 21 24 27

Using the Equation 1 we find that in this ordered array the median will be situated between 4th and 5th scores (Equation 7):

$$Me_{Rank} = \frac{(8+1)}{2} = 4.5 \qquad \textbf{Equation 7.}$$

So, the median is equal to 19.5.

Example A - calculation of the median in a groupped data

In a study of pain measurement the following data were obtained for 17 units (n=17) (2):

1 1 2 2 2 2 2 3 3 3 3 4 4 4 5 5 5

By inspection, we can see that the median will fall into the category "3" (the 9th score).
Assuming underlying continuity, we can present the data into real class intervals (Table 2):

Table 2. Data on a measurement of pain among 17 subjects

Score	Real class interval	Frequency	Cumulative frequency
1	0.5 – 1.4	2	2
2	1.5 – 2.4	5	7
3	2.5 – 3.4	4	11*
4	3.5 – 4.4	3	14
5	4.5 – 5.4	3	17
		N = 17	* median

Now we can apply Equation 2 for calculation (Equation 8):

$$Me = 2.5 + \left(1 \times \frac{(17/2) - 7}{4}\right) = 2.5 + \left(\frac{1.5}{4}\right) = 2.5 + 0.375 = 2.875 \qquad \textbf{Equation 8.}$$

Example 3 – calculation of percentiles in an ungroupped data

Let's say that a student received a score of 90 points on a test given to a group of 50 examined students, 40 of them with scores less than 90 points (2). That means that his location in an ordered series is 41st place. The percentile rank for this student will be calculated according to Equation 3 (Equation 9):

$$P_i = \frac{40}{50} \times 100 = 80 \qquad \textbf{Equation 9.}$$

In other words, he achieved a higher score than 80% of the students who took the test and 20% of the students received better results than that particular student.

Example 3 – calculation of percentiles in an inteval scale and its use

In a representative sample of 120 urban male liveborns the height scores are presented in an interval array of 3 cm of width. Let's calculate the value of P_3, P_{10}, P_{25}, P_{50}, P_{75}, P_{90} and P_{97}. The values of the variable "weight" are presented in a table (Table 3), containing equal width intervals and their corresponding frequencies.

The calculation of percentiles includes the following steps (3,7,8):

1. Determining the cumulative frequences for each interval by adding up to the absolute frequences in a given interval the absolute frequences from the previous interval.
2. Determing the percentile ranks using the Equation 3 (Equations 10-16):

Table 3. The values of the variable "weight" containing equal width intervals and their corresponding frequencies

Height in cm *x*	Frequency f	Cumulative frequency f	Percentiles' rank
41 - 43	3	-	
44 - 46	14	17	$P_{3,}\ P_{10}$
47 - 49	40	57	P_{25}
50 - 52	45	102	$P_{50,}\ P_{75}$
53 - 55	15	117	$P_{90,}\ P_{97}$
56 - 58	3	120	
	$\Sigma f = N = 120$		

$$P_3 = \frac{120}{100} \times 3 = 3.6 \qquad \textbf{Equation 10.}$$

$$P_{10} = \frac{120}{100} \times 10 = 12 \qquad \textbf{Equation 11.}$$

$$P_{25} = \frac{120}{100} \times 25 = 30 \qquad \textbf{Equation 12.}$$

$$P_{50} = \frac{120}{100} \times 50 = 60 \qquad \textbf{Equation 13.}$$

$$P_{75} = \frac{120}{100} \times 75 = 90 \qquad \textbf{Equation 14.}$$

$$P_{90} = \frac{120}{100} \times 90 = 108 \qquad \textbf{Equation 15.}$$

$$P_{97} = \frac{120}{100} \times 97 = 116.4 \qquad \textbf{Equation 16.}$$

3. According to the cumulative frequencies and the percentile ranks in step 2 we define for each percentile the interval where it should fall: in our example P_3 and P_{10} fall in the interval 44-46 cm, P_{25} - in the interval 47-49, P_{50} and P_{75} - in the interval 50-52 cm, P_{90} and P_{97} - in the interval 53-55 cm.
4. Using the Equation 2, we can calculate the value of each percentile. Here we need to take into account that value of "c" - the difference between the percentile rank and the cumulative frequency in the previous interval is:
   - for P_3 = 3.6 (percentile rank) – 0 (cumulative frequency in the previous interval) = 3.6,
   - for P_{10} = 10 (percentile rank) – 0 (cumulative frequency in the previous interval) = 10,
   - for P_{25} = 30 (percentile rank) – 17 (cumulative frequency in the previous interval)= 13,
   - for P_{50} = 60 (percentile rank) – 57 (cumulative frequency in the previous interval) = 3,
   - for P_{75} = 90 (percentile rank) – 57 (cumulative frequency in the previous interval) = 33,
   - for P_{90} = 108 (percentile rank) – 102 (cumulative frequency in the previous interval) = 6,
   - for P_{97} = 116.4 (percentile rank) – 102 (cumulative frequency in the previous interval) = 14.4.

   The values for each percentile are now calculated as follows (Equations 17-23):

$$P_3 = 44 + \frac{3.6}{14} \times 3 = 44.77 \qquad \textbf{Equation 17.}$$

$$P_{10} = 44 + \frac{10}{14} \times 3 = 46.14 \qquad \textbf{Equation 18.}$$

$$P_{25} = 47 + \frac{13}{40} \times 3 = 47.98 \qquad \textbf{Equation 19.}$$

$$P_{50} = 50 + \frac{3}{45} \times 3 = 50.20$$ **Equation 20.**

$$P_{75} = 50 + \frac{33}{45} \times 3 = 52.20$$ **Equation 21.**

$$P_{90} = 53 + \frac{6}{15} \times 3 = 54.20$$ **Equation 22.**

$$P_{97} = 53 + \frac{14.4}{15} \times 3 = 55.88$$ **Equation 23.**

We can use these data now to determine the reference groups for the weight of the male newborns in that urban area. The limits of the seven reference groups will be as presented in Table 4.

Table 4. Percentiles and their use for reference groups*.

Reference groups	Percentiles	Weight of the male newborns
Severely stunted	Below P_3	Below 44.77
Moderately below the norm	$P_3 - P_{10}$	44.77 – 46.14
Sligthly below the norm	$P_{10} - P_{25}$	46.14 – 47.98
Normal range	P_{25} - P_{75}	47.98 – 52.2
Sligthly above the norm	$P_{75} - P_{90}$	52.2 – 54.2
Moderately above the norm	$P_{90} - P_{97}$	54.2 – 55.88
Expressed accelaration	Above P_{97}	Above 55.88

* All the data in the table are an approximation

Example 5 – identification of quartiles in an ungroupped data

Say we have a sample where n = 16 and the values of the variable are the following (2):

1 5 7 7 8 9 9 10 11 12 13 15 19 20 20 20

By inspection of the data, we find:

- the first quartile Q_1 is located between 4^{th} and 5^{th} scores; thus $Q_1 = (7 + 8)/2 = 7.5$
- the second quartile Q_2 is located between the 8^{th} and 9^{th} scores; thus $Q_2 = (10 + 11)/2 = 10.5$
- the third quartile Q_3 is located between 13^{th} and 14^{th} scores; thus $Q_3 = (15 + 19)/2 = 17.0$

These values are located in the array as follows:

1 5 7 [7 8] 9 9 [10 11] 12 13 [15 19] 20 20 20

Exercises

Task 1

For the following questions choose between "true" or "false":

1. The 50^{th} percentile score and the median will always be the same value. T F
2. Twenty five percent (25%) of the scores fall between Q1 and the median. T F
3. The distance between Q1 and the median is always different to the distance between Q3 and the median. T F
4. The median and the 50^{th} percentile rank have different values. T F

Answers: 1. T; 2. T; 3. F; 4. F

Task 2

Consider following set of data:

3, 3, 4, 5, 6, 7, 8, 9, 9, 10, 38, 60

1. The median is:
   A. 7.0
   B. 7.5

C. 8.0
D. 3 or 9

2. Q_1 is:
A. 4.5
B. 5.5
C. 8.0
D. 9.5

3. Q_3 is:
A. 4.5
B. 6.0
C. 7.5
D. 9.5

Answers: 1. B; 2. A, 3. D

References

1. Last JM (editor). A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Polgar S, Thomas SA. Introduction to research in the health sciences. 2nd ed. Singapore: Churchill Livingstone; 1991.
3. Grancharova G, Hristova P. Medical statistics (in Bulgarian). Pleven: Medical University of Pleven Publishing House, 2004.
4. Leaverton PE. A review of biostatistics. A program for self-instruction. 2nd edition. Boston: Little Brown and Company; 1981.
5. Munro BH, Visintainer MA, Page EB. Statistical methods for health care research. London: J. B. Lippincott Company; 1986.
6. Lwanga SK, Tye CY, Ayeni O. Teaching health statistics. Lessons and seminar outlines. 2nd ed. Geneva: World health organization; 1999.
7. Kalinov K. Statistical methods in behavioural and social sciences (in Bulgarian). Sofia: New Bulgarian University Publishing House; 2001.
8. Ranchov G. Biostatistics and Biomathematics – concepts, methods, applications (in Bulgarian). Sofia: Eco Print; 2008.
9. Palmer C. Lectures in biostatistics. Teaching Course in Modern Epidemiology. Varna: Medical University of Varna; 1994.

Recommended readings

1. Campbell MJ, Machin D, Walters SJ. Medical Statistics: A textbook for the health sciences. 4th rev. ed. Chichester: John Wiley & Sons LTD; 2007.
2. Kirkwood BR, Sterne JAC. Essentials of medical statistics. 2nd rev. ed. Chichester: Blackwell Publishing Company, 2003.
3. Polgar S, Thomas SA. Introduction to research in the health sciences. 2nd ed. Singapore: Churchill Livingstone; 1991.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Frequency measures: estimating risk**
Module: 2.7	ECTS (suggested): 0.3
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Jadranka Bozikov, dipl.ing.math, PhD, Professor Andrija Stampar School of Public Health, Medical School, University of Zagreb, Croatia
Address for correspondence	Lijana Zaletel-Kragelj Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Actuarial method, cumulative incidence, density method, risk, simple cumulative method.
Learning objectives	After completing this module students should be: • familiar with differences between four different methods for estimation of cumulative risk, being simple cumulative, actuarial, density, and Kaplan Meier method; • able to estimate cumulative risk measures of different levels of accuracy independently.
Abstract	Risk is defined as the probability that a disease-free individual is developing a disease under observation over a specified period, conditional on that the same individual is not dying from any other disease during the period. In practice, risk is estimated by using different methods. The simple cumulative method is the easiest and most widely used. Risk cannot be accurately estimated by this method unless all subjects in the observed candidate population are followed for the entire follow-up period, or are known to develop the disease during the period (no censoring). Because of serious limitations of this method, several methods more or less susceptible to censoring were proposed. Considering the censoring of the data in estimating cumulative risk requires the use of special analytic methods. These methods are actuarial, density, and Kaplan Meier method.
Teaching methods	An introductory lecture gives the students first insight into four methods for calculation of cumulative risk. The theoretical knowledge is illustrated by case studies. After introductory lectures, students first carefully read the theoretical background of this module and complement their knowledge with recommended readings. Afterwards, they perform tasks on a provided data set on estimation of different types of measures. Students are stimulated to compare results with their peers and discuss potential differences.
Specific recommendations for teachers	• work under teacher supervision/individual work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on calculated measures in which a detailed description of the process of calculation is described.

FREQUENCY MEASURES: ESTIMATING RISK

Lijana Zaletel-Kragelj, Jadranka Bozikov

Theoretical background

Introduction

In expressing relative incidence we are dealing with several measures. One of them is the so-called risk.

Risk is defined as the probability that a disease-free individual is developing a disease under observation over a specified period, conditional on that the same individual is not dying from any other disease during the period (1). Thus, risk is a conditional probability, with values varying between zero and one. It is dimensionless (1). It usually refers to the first occurrence of the disease for each initially disease-free individual, although it is possible to consider the risk of developing the disease under observation within a specified period more than once (1).

In practice, risk is estimated by using different methods. The simple cumulative method is the easiest and the most widely used (1). For a cohort of subjects followed for a given period of time, risk is often estimated by calculating the proportion of candidate subjects who develop the disease during the observation period. This measure is usually referred to as the cumulative incidence (CI) (1). Generally, cumulative incidence is estimated only for the first occurrence of the disease. If the durations of the individual follow-up periods for all non-cases are equal, the cumulative incidence is equivalent to the average risk for members of the cohort. This means that under the condition of a fixed cohort, cumulative incidence is a good estimate of risk. This is the reason that cumulative incidence and risk are frequently equalized. But once again, because risk is, by its definition, a conditional probability, it cannot be accurately estimated by calculating cumulative incidence unless all subjects in the observed candidate population are followed for the entire period, or are known to develop the disease (or other observed phenomenon) during the period (1).

The cumulative probability of the event during a given time interval is the proportion of new events during the interval in which the denominator is the initial number of observed persons. The calculation of this measure is straightforward if no losses happen in the cohort during the interval (1-9). However, in real life the size of the cohort is more than likely to be decreased after a long period of follow-up as a result of different reasons. A situation in which the event and the time of individual at risk for the event is unknown is usually called censoring (2,8-12).

There are usually three reasons why censoring occurs. The first is the termination of the observation because of the end of the study before the event occurs, the second is the termination because of some competing factors (death of other cause e.g. traffic accident), the third is simply the lost because of changing the domicile of the individual under observation, etc. In all cases, the occurrence of observed phenomenon is unknown. The terms also used with this phenomenon are “withdrawals”, “losts-to-follow-up” and others (2,8-12). Considering the censoring of the data requires the use of special analytic methods. The methods of risk estimation are the simple cumulative method, the actuarial method, the density method, and the Kaplan Meier product limit method (1,2,9-13).

Methods of risk estimation

Simple cumulative method

This method is the easiest for estimating risk (1,2,12). The risk calculated by this method is the roughest estimate in this family of measures.

It is simply the proportion of new events during the interval in which the denominator is the initial number of observed persons (Equation 1):

$$ {}_{cum}R = \frac{N_{d+new\,cases(gp)}}{N_{all\,persons\,at\,risk(bgp)}} \qquad \textbf{Equation 1.} $$

${}_{cum}R$ = cumulative risk (risk of getting a disease during the entire period)
$N_{d+\ new\ cases\ (gp)}$ = number of new cases of the disease under observation during a given period
$N_{all\ persons\ at\ risk\ (bgp)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given period

Usually it is estimated only for the first occurrence of the disease. This is the reason that the population at risk (the denominator in the equation) consists of disease-free individuals at the beginning of the observational period. The observation period has to be clearly stated since the value of the measure is increasing with the prolongation of period of observation. This period could be based upon a calendar time or not (e.g. first year after the exposure, first year after surgery etc.). It is a good estimate of the risk only in the case of fixed cohorts in which there are no withdrawals from the follow-up (1,12).

Estimation of cumulative risk over the entire 5-year observational period in practice is presented in Case study 1.

For avoiding the drawbacks of this rough direct method of estimation of cumulative risk over longer periods, we could split this longer period first to shorter periods (i.e. 1-year periods) and obtain cumulative risk indirectly

through calculating risks for these periods (partial risks). When partial risk refers to 1-year period it is known as annual risk (Equation 2):

$$_{ann}R = \frac{N_{d+ new cases (1\text{-}year\ period)}}{N_{all\ persons at risk (beginning of\ 1\text{-}year\ period)}} \qquad \textbf{Equation 2.}$$

$_{ann}R$ = annual risk (risk of getting a disease during the 1-year period)
$N_{d+ new cases (1\text{-}year period)}$ = number of new cases of the disease under observation during1-year period
$N_{all persons at risk (beginning of 1\text{-}year period)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given 1-year period

The annual risk is annual probability of the event (12). The complement of this probability (the mirror image) is annual probability of survival without an event under observation (i.e. a breakout of a disease). Technically, these probabilities are conditional probabilities. This means for example, that one has to survive through the first interval in order to be a part of the denominator for the calculation of the survival probability in the second interval. Similarly, the survival probability for the third interval is calculated only among those persons who survived in the first and then in the second interval (12).

A cumulative probability of survival without a disease under observation over more than one interval (2-, 3-, 4-, 5-year interval, etc.) is obtained by multiplying the annual conditional survival probabilities over all intervals (12). Afterwards, we calculate again complementary values (1 – cumulative survival) that are in fact cumulative risks over more than one interval.

By using this procedure the censoring is partially considered even when using simple methods, as we need to define separately for every year the number of individuals under observation at risk, and all participants who terminated the observation because of extraneous factors (e.g. death because of traffic accidents, etc.) are not included.

Estimation of cumulative 5-year risk over observational period through calculation of annual risks is presented in Case study 1.

Actuarial method

This is the first method in which the censoring is considered in the calculation of risk estimate (1,8,11-13). It is typically used to estimate the probability of death in survival analysis, but as mortality is a special case of incidence (12), it could be generalized to estimation of risk on general (2). It is referred also as interval-based life table or life table interval approach (12).

This method is working under the assumption that the censoring is occurring uniformly throughout the observed period (usually meaning that all withdrawals, i.e. censored observations, occur on average in the middle of the observational period) (1,2,11). If the periods are short (up to 1 year), or there is a small number of withdrawals this assumption does not affect the risk estimate seriously (1). However, one should be aware that this method still provides us more or less biased estimates of risk (1). The basic equation for calculating risk by using actuarial method directly is as follows (Equation 3):

$$_{cum}R = \frac{N_{d+ new cases (gp)}}{N_{all\ persons at risk (bgp)} - \frac{N_{w(gp)}}{2}} \qquad \textbf{Equation 3.}$$

$_{cum}R$ = cumulative risk (risk of getting a disease during the entire period)
$N_{d+ new cases (gp)}$ = number of new cases of the disease under observation during a given period
$N_{all persons at risk (bgp)}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of a given period
$N_{w (gp)}$ = number of withdrawals during a given period

For avoiding the drawbacks of this method we could again split longer periods first to shorter periods (i.e. 1-year periods) and calculate risks for these periods (i.e. annual risks). Only afterwards, on the basis of risks of shorter periods as intermediate elements, the cumulative risk is calculated indirectly. Annual risks could be calculated as follows (Equation 4):

Estimation of this measure in practice is presented in Case study 2.

Again, a cumulative probability of survival without a disease under observation over more than one interval (2-, 3, 4-, 5-year interval, etc.) is obtained by multiplying the partial conditional survival probabilities over all intervals (Equation 5) (12):

$$_{ann}R = \frac{N_{d+ new cases (1\text{-}year\ period)}}{N_{all\ persons at risk (beginning of\ 1\text{-}year\ period)} - \frac{N_{w(1\text{-}year\ period)}}{2}} \qquad \textbf{Equation 4.}$$

$_{ann}R$ = annual risk (risk of getting a disease during the 1-year period)
$N_{d+ new cases (1\text{-}year period)}$ = number of new cases of the disease under observation during the 1-year period

$N_{\text{all persons at risk (beginning of 1-year period)}}$ = number of all persons at risk for getting ill with the disease under observation at the beginning of the 1-year period

$N_{\text{w (1-year period)}}$ = number of withdrawals during the 1-year period

$$_{cum}R = 1 - \left[\left(1 - {}_{ann}R_{(year\ 1)}\right) \times \left(1 - {}_{ann}R_{(year\ 2)}\right) \times \ldots \times \left(1 - {}_{ann}R_{(year\ n)}\right)\right] \quad \textbf{Equation 5.}$$

$_{cum}R$ = cumulative risk (risk of getting a disease during the entire period of observation)

$_{ann}R_{(year\ 1)}$ = annual risk (risk of getting a disease) during the 1st year

$_{ann}R_{(year\ 2)}$ = annual risk (risk of getting a disease) during the 2nd year

$_{ann}R_{(year\ n)}$ = annual risk (risk of getting a disease) during the nth year

Estimation of this measure in practice is presented in Case study 2.

Because of serious limitations of this method, other methods were proposed (1).

Density method

Actuarial method is working under the assumption that all withdrawals occur on average in the middle of the observational period (1,2,11). If the periods are short, or there is a small number of withdrawals this assumption does not affect the risk estimate seriously (1). However, it is better to consider exact times of being at risk of developing a disease under observation. Another interval-based method based on the estimation of average incidence rates (person-time rate or incidence density) was proposed (1,3,4,11,12). This method depends on the functional relationship between a risk and an incidence rate (estimated through incidence density) (1).

Risk depends on incidence density and on the duration of the period of observation. Under the assumption that the cohort under observation is fixed (with no censored observations), and that the incidence density is constant over the period of observation, the risk estimate could be directly calculated as follows (Equation 6) (1,3):

$$_{cum}R = 1 - e^{\left(-ID \times t_{(gp)}\right)} \quad \textbf{Equation 6.}$$

$_{cum}R$ = cumulative risk (risk of getting a disease during the entire period)

ID = incidence density

$t_{(gp)}$ = duration of the given period of observation (period at risk)

Incidence density, used in this equation was introduced in a separate module in this book. It is the rate between the number of new cases which occur during the period under observation, and the quantity known under the term person-time (PT). It is calculated as follows (Equation 7):

$$ID = \frac{N_{d+\ new\ cases\ (gp)}}{PT} \quad \textbf{Equation 7.}$$

ID = incidence density

$N_{d+\ new\ cases\ (gp)}$ = number of new cases of the disease under observation during a given period

PT = person-time

However, usually the incidence density (as an estimate of incidence rate) does not remain constant during the entire follow-up period. Like in the actuarial method, cumulative risk over a longer period also in this method is not calculated directly. We split this longer period first to shorter periods (i.e. 1-year periods) and calculate risks for these periods (partial risks), i.e. annual risks. These could be calculated as follows (Equation 8):

$$_{ann}R = 1 - e^{\left(-{}_{ann}ID \times 1\right)} \quad \textbf{Equation 8.}$$

$_{ann}R$ = annual risk (risk of getting a disease during the 1-year period)

$_{ann}ID$ = annual incidence density

We can see that annual incidence densities need to be calculated prior to calculation of annual risks (Equation 9):

$$_{ann}ID = \frac{N_{d+\ new\ cases\ (1\text{-}year\ period)}}{PT} \quad \textbf{Equation 9.}$$

$_{ann}ID$ = annual incidence density

$N_{d+\ new\ cases\ (gp)}$ = number of new cases of the disease under observation during a 1-year period

PT = person-time

Estimation of annual incidence densities and annual risks estimated by using the density method in practice is presented in Case study 3.

Only afterwards, on the basis of annual risks as intermediate elements, the cumulative risk is calculated as follows (Equation 10):

$$_{cum}R = 1 - e^{\left[\left(-{}_{ann}ID_{(year\ 1)} \times 1\right) + \left(-{}_{ann}ID_{(year\ 2)} \times 1\right) + \ldots + \left(-{}_{ann}ID_{(year\ n)} \times 1\right)\right]} \quad \textbf{Equation 10.}$$

$_{cum}R$ = cumulative risk (risk of getting a disease during the entire period)

$_{ann}ID_{(year\ 1)}$ = annual incidence density during the 1st year

$_{ann}ID_{(year\ 2)}$ = annual incidence density during the 2nd year

$_{ann}ID_{(year\ n)}$ = annual incidence density during the nth year

Estimation of this measure in practice is presented in Case study 3.

Kaplan Meier product limit method

Kaplan Meier product limit method (8,11,12) combines calculated probabilities of survival and estimates to allow censored observations, which are assumed to occur randomly. The intervals are defined as ending each time an event (i.e. disease, death, withdrawal) occurs and are, therefore, unequal (2,12). Again, these probabilities are conditional – they are conditioned on being at risk (present in the study without a disease under observation or censored) at each event time. The formula for calculation of conditional probability is as follows (Equation 11):

$$p = \frac{N_{d+i}}{N_{persons\,at\,risk\,i}} \qquad \textbf{Equation 11.}$$

p = conditional probability for an event in time i
N_{d+i} = number of events (new cases of a disease or death) occurring at time i
$N_{persons\ at\ risk\ i}$ = number of individuals still under observation (still at risk of the event under observation) at time i

When time i is measured exactly, the number of events is usually 1.

The complement of this conditional probability of an event is probability of survival without an event under observation (i.e. a breakout of a disease) (12). A cumulative probability of survival without a disease under observation over more than one interval (2-, 3-, 4-, 5-year interval, etc.) is obtained by multiplying the annual conditional survival probabilities over all intervals (12).

Estimation of cumulative 5-year risk over observational period through calculation of conditional probabilities is presented in Case study 4.

Case studies

Data set

For the illustration of differences between the simple, the actuarial, the density, and the Kaplan Meier product limit method of calculation of cumulative risk an imaginary data-set is used. A cohort of 20 individuals initially without a disease under observation, were followed up for 5 years (Figure 1).

In this period, 16 individuals got a disease under observation (an event under observation) (Figure 1, persons with black lines of follow-up time), while 4 of them were lost to follow-up because of voluntarily withdrawal from the study or change of domicile (persons No. 5, 7, 14 and 19) (Figure 1, persons with grey lines of follow-up time). The lines with arrows indicate that individuals were alive at the time of the lost of follow-up.

In Figure 1 the members of a cohort are presented in order as they were numbered at the time of the entry into the study, while in Figure 2, the members are rearranged in rank order regarding the time of an event or withdrawal. This presentation is useful in determination of times of being at risk for the event under observation.

Case study 1: Estimation of cumulative risk using simple cumulative method

Results of counting of cases of the observed disease which broke out during the entire 5-year time of observation (Figure 1) show that the cumulative 5-year risk estimated by the simple cumulative method according to Equation 1 is as follows (Equation 12):

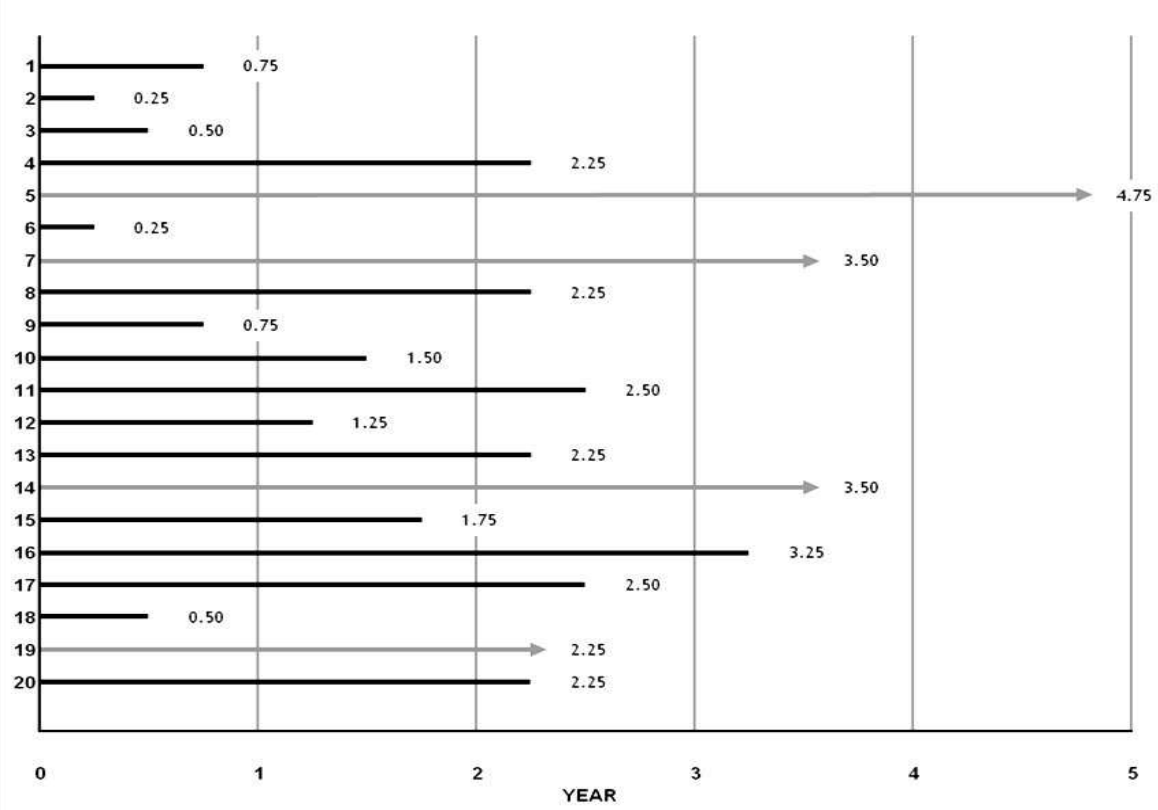
$$_{cum}R = \frac{16}{20} = 0.8000 \qquad \textbf{Equation 12.}$$

But, this estimate is unreliable as there are censorings in 4/20 individuals under observation (No. 5, 7, 14 and 19) (Figure 1). In these individuals the occurrence of the event of interest is uncertain because of the termination of the observation before the event occurred. To diminish the drawbacks of this method we can split the 5-year interval into 5 five 1-year intervals, and for each 1-year interval we calculate the annual risk by following the next steps:

- define the number of persons entered in the interval (Table 1, column 1), number of persons with the disease at the end of interval (Table 1, column 2), and the number of losts (withdrawals) (Table 1, column 3),
- by using Equation 2 calculate annual risks (Table 1, column 4).

From the Table 1 it could be seen that in case of calculation of annual risks, the censoring is partially considered even when using the simple cumulative method, as we need to define separately for every year the number of individuals at risk, and all participants who terminated the observation because of extraneous factors (e.g. death because of traffic accidents, etc.) are not included.

Figure 1. Graphic presentation of events in a cohort of 20 people



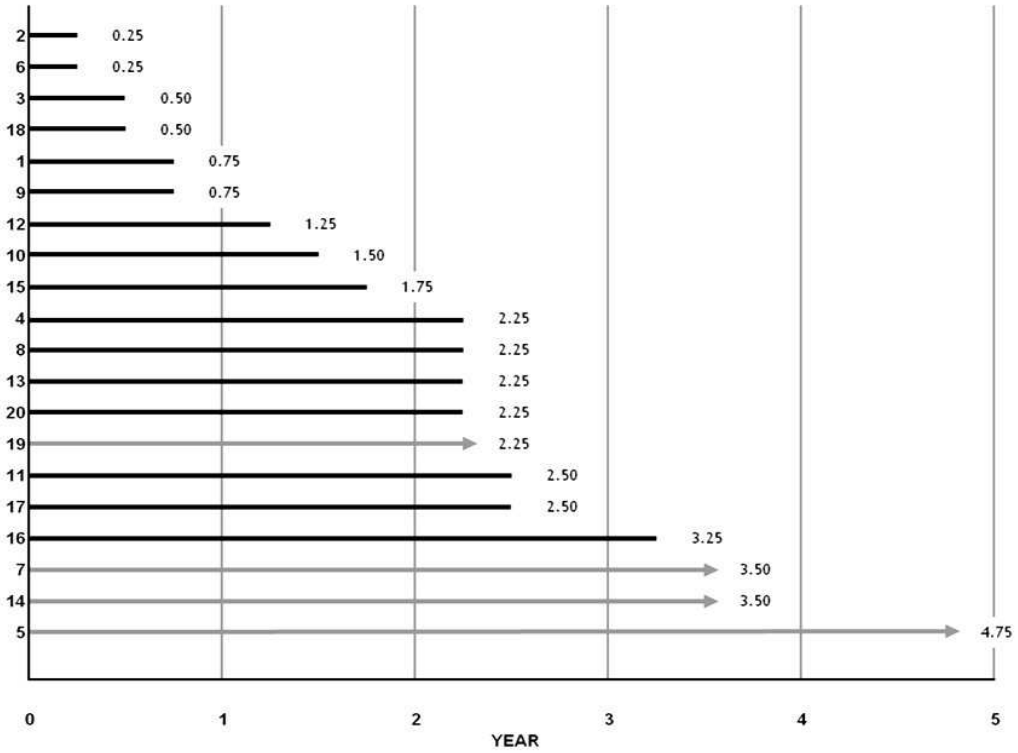
LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Table 1. Elements for calculation and calculation of annual risks using the simple cumulative method

Year of observation	1	2	3	4
	Entered in the interval (N)	With the disease at the end of interval (d+)	Lost	d+/N (annual risk) ($_{ann}$R)
1st	20	6	0	0.3000
2nd	14	3	0	0.2143
3rd	11	6	1	0.5455
4th	4	1	2	0.2500
5th	1	0	1	0.0000

The annual risk (Table 1, column 4) is annual probability of the event (12). The complement of this probability is annual probability of survival without an event under observation (i.e. a breakout of a disease) (Table 2, column 5). Technically these probabilities are annual conditional probabilities. A cumulative probability of survival without a disease under observation over more than one interval (2-, 3-, 4-, and 5-year interval) is obtained by multiplying the annual conditional survival probabilities over all intervals (Table 2, column 6) (12).

Figure 2. Ordered time of being at risk of developing a disease under observation in a cohort of 20 people from Figure 1



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Table 2. Calculation of cumulative 5-year risk from annual risks using simple cumulative method.

Year of observation	4	5	6	7
	d+/N (annual risk) ($_{ann}$R)	1 - $_{ann}$R	product (1 - $_{ann}$R) (Π)	1 – Π (cumulative risk) ($_{cum}$R)
1st	0.3000	0.7000	0.7000	0.3000
2nd	0.2143	0.7857	0.5500	0.4500
3rd	0.5455	0.4545	0.2500	0.7500
4th	0.2500	0.7500	0.1875	0.8125
5th	0.0000	1.0000	0.1875	0.8125

The cumulative probability of having an event is the complement of the joint probability of survival through each of five years of observation (Table 2, column 7) (12).

Case study 2: Estimation of cumulative risk using actuarial method

Simple cumulative method assumes no withdrawals during the period of observation. Since in our case (Figures 1 and 2) there were four individuals lost to observation, this must be considered. Their limited participation need to be considered in the denominator of the cumulative probability of an event. Actuarial method considers censored observations most roughly (Equation 3). Since we have at the end of the 5-year interval 16 individuals with a disease out of 20 persons at the beginning of the observation, and 4 persons were lost to follow up, we calculate cumulative 5-year risk directly as follows (Equation 13):

$$_{cum}R = \frac{16}{20 - \frac{4}{2}} = 0.8889 \qquad \textbf{Equation 13.}$$

Again, we can split 5-year interval first into five 1-year intervals and calculate first the annual risks and afterwards cumulative 5-year risk. For each 1-year interval we:

- define the number of persons entered in the interval (Table 3, column 1), number of persons with the disease at the end of interval (Table 3, column 2), and the number of withdrawals (Table 3, column 3),
- calculate the adjusted number of withdrawals (1,12),
- by using Equation 4 calculate annual actuarial risks (Table 3, column 6).

After annual risks are calculated we follow exactly the same principles for calculation of 2-, 3-, 4- and 5-year cumulative risks as discussed in the simple method. The results are presented in Table 4. Results of calculating the cumulative 5-year risk estimated by using the actuarial method (Table 4, column 9) show that its value is 0.8428, which is much higher than estimated by using the simple method.

Table 3. Elements for calculation and calculation of annual risks using actuarial method

Year of observation	1	2	3	4	5	6
	Entered in the interval (N)	With the disease at the end of interval (d+)	Withdrawals (W)	W/2	N – (W/2)	d+/N–(W/2) (annual risk) ($_{ann}R$)
1st	20	6	0	0	20	0.3000
2nd	14	3	0	0	14	0.2143
3rd	11	6	1	0.5	10.5	0.5714
4th	4	1	2	1	3	0.3333
5th	1	0	1	0.5	0.5	0.0000

Table 4. Calculation of cumulative 5-year risk from annual risks using actuarial method

Year of observation	6	7	8	9
	d+/N – (W/2) (annual risk) ($_{ann}R$)	1 – $_{ann}R$	product (1 - $_{ann}R$) (Π)	1 – Π (cumulative risk) ($_{cum}R$)
1st	0.3000	0.7000	0.7000	0.3000
2nd	0.2143	0.7857	0.5500	0.4500
3rd	0.5714	0.4286	0.2357	0.7643
4th	0.3333	0.6667	0.1572	0.8428
5th	0.0000	1.0000	0.1572	0.8428

Case study 3: Estimation of cumulative risk using the density method

The first method that considers exact times of being at risk of developing a disease under observation is the density method.

In order to perform the procedure (Equation 6) we need first to calculate the person-years (PY) since we need this quantity in calculation of the incidence density. We use the information given in Figure 2. In Table 5, data for calculation of PY for the entire 5-year period are presented. The incidence density for 5-year period could be now calculated using the Equation 7. The results are presented in the following equation (Equation 14):

$$\mathrm{ID} = \frac{16}{38.75} = 0.4129 \qquad \textbf{Equation 14.}$$

This quantity afterwards enters the equation for calculating the 5-year cumulative risk using the Equation 6. The results are presented in the following equation (Equation 15):

$$_{\mathrm{cum}}\mathrm{R} = 1 - \mathrm{e}^{(-0.4129 \times 5)} = 0.8731 \qquad \textbf{Equation 15.}$$

Table 5. Data for calculation of person-years

Id. number	Time of being at risk* (Years)	Status at the end of observation (1=with the disease, 0=censored (cause of censoring))
2	0.25	1
6	0.25	1
3	0.50	1
18	0.50	1
1	0.75	1
9	0.75	1
12	1.25	1
10	1.50	1
15	1.75	1
4	2.25	1
8	2.25	1
13	2.25	1
20	2.25	1
19	2.25	0 – free of disease, change of domicile
11	2.50	1
17	2.50	1
16	3.25	1
7	3.50	0 – free of disease, voluntarily withdrawal
14	3.50	0 – free of disease, change of domicile
5	4.75	0 – free of disease, change of domicile
Total	38.75	Diseased = 16, Lost-to-follow-up = 4

* time in which an individual under observation is exposed to effect of noxious agent (is at risk of getting an event under observation)

Again, we can split 5-year interval first into five 1-year intervals and calculate initially the annual risk using the density method and afterwards the cumulative 5-year risk. The steps are as follows

- first, we summarize the events in each of 1-year intervals which are five as the duration of the longest observation is 4.75 let: entered in the interval (Table 6, column 1), with the disease at the end of interval (Table 6, column 2), lost to follow-up (Table 6, column 3), and present at the end of the period without disease (Table 6, column 4),

Table 6. Summary of the events in each of 1-year intervals

Year of observation	1	2	3	4
	Entered in the interval (N)	With the disease at the end of interval (d+)	Lost to follow-up	Present at the end of the period
1st	20	6	0	14
2nd	14	3	0	11
3rd	11	6	1	4
4th	4	1	2	1
5th	1	0	1	0
Total		16	4	

- in the following step we calculate the person-years (PY) for each of 1-year periods (Table 7). We use the information given in Figure 2:

Table 7. The summary of calculation of person-years in each of five 1-year intervals

Year of observation	Contribution to person-years (PY) at the end of 1-year interval	PY Total
1st	(0.25 × 2) + (0.50 × 2) + (0.75 × 2) + (1.00 × 14)	17.00
2nd	(0.25 × 1) + (0.50 × 1) + (0.75 × 1) + (1.00 × 11)	12.50
3rd	(0.25 × 5) + (0.50 × 2) + (1.00 × 4)	6.25
4th	(0.25 × 1) + (0.50 × 2) + (1.00 × 1)	2.25
5th	(0.75 × 1)	0.75

- in the following step the annual incidence density is calculated (Table 8). As the incidence density is not constant over the 5-year period (the highest is in the third year of observation) this has to be considered in the calculation of the cumulative risk.
- as a final step, the risk is calculated based on the incidence density (Table 9).

Table 8. Calculation of incidence density in each of five 1-year intervals

Year of observation	2	5	6
	With the disease at the end of interval (d+)	Annual person-years (PY)	Annual incidence density (d+/PY) ($_{ann}$ID)
1st	6	17.00	0.3529
2nd	3	12.50	0.2400
3rd	6	6.25	0.9600
4th	1	2.25	0.4444
5th	0	0.75	0.0000

Results of calculating the cumulative 5-year risk estimated by using the density method (Figure 1) show that its value is 0.8643, which is much higher than estimated using the simple method, and also higher than estimated using the actuarial method. The elements for calculation and its results are presented in Table 10.

Table 9. Calculation of the annual risk in each of five 1-year intervals

Year of observation	6	7	8
	Annual incidence density (d+/PY) ($_{ann}$ID)	$e^{(-IDann\times1)}$	$1-e^{(-annID\times1)}$ (annual risk) ($_{ann}$R)
1st	0.3529	0,.7027	0.2974
2nd	0.2400	0.7866	0.2134
3rd	0.9600	0.3829	0.6171
4th	0.4444	0.6412	0.3588
5th	0.0000	1.0000	0.0000

Table 10. Elements for calculation of cumulative risk using the density method

	6	9	10	11
Year of observation	Annual incidence density (d+/PY) ($_{ann}$ID)	$\sum(-_{ann}ID\times1)$	$e^{\sum(-annID\times1)}$	$1-e^{\sum(-annID\times1)}$ (cumulative risk) ($_{cum}$R)
1st	0.3529	-0.3529	0.7026	0.2974
2nd	0.2400	-0.5929	0.5527	0.4473
3rd	0.9600	-1.5529	0.2116	0.7884
4th	0.4444	-1.9973	0.1357	0.8643
5th	0.0000	-1.9973	0.1357	0.8643

Case study 4: Estimation of cumulative risk using Kaplan-Meier product limit method

This method also considers exact times of being at risk of developing a disease under observation (2,12). The intervals are defined as ending each time an event (i.e. disease, death, withdrawal) occurs. The procedure is as follows:

- first we determine the times when events or censoring occurred. We use the information given in Figure 2,
- define the number of persons entered in the interval (Table 11, column 1), number of persons with the event (occurrence of the disease or death) at time I (Table 11, column 2), and the number of censored cases (Table 11, column 3) at time i,
- by using Equation 11 calculate conditional probabilities (Table 11, column 4),
- calculate the complement of conditional probabilities of the event at every time of occurrence of the events or censoring – the conditional probability of survival without an event under observation up to the time i (Table 11, column 5),

Table 11. Elements for calculation of cumulative risk by using the Kaplan-Meier product limit method

Time of the events/ censoring (years)	1	2	3	4	5	6	7
	Entered in the interval (N)	Occurrence of the event (d+)	Censored	d+/N (conditional probability of the event) (p)	1 – p (survival) (S)	Product (S) (Cumulative survival) ($_{cum}$S)	1 – S_{cum} (cumulative conditional probability of an event) ($_{cum}$R)
0.25	20	2	0	0.1000	0.9000	0.9000	0.1000
0.50	18	2	0	0.1111	0.8889	0.8000	0.2000
0.75	16	2	0	0.1250	0.8750	0.7000	0.3000
1.25	14	1	0	0.0714	0.9286	0.6500	0.3500
1.50	13	1	0	0.0769	0.9231	0.6000	0.4000
1.75	12	1	0	0.0833	0.9167	0.5500	0.4500
2.25	11	4	1	0.3636	0.6364	0.3500	0.6500
2.50	6	2	0	0.3333	0.6667	0.2333	0.7667
3.25	4	1	0	0.2500	0.7500	0.1750	0.8250
3.50	2	0	2	0.0000	1.0000	0.1750	0.8250
4.75	1	0	1	0.0000	1.0000	0.1750	0.8250

- calculate cumulative probability of survival over more than one interval by multiplying the conditional survival probabilities over all intervals (Table 11, column 6),
- calculate the complement of cumulative probabilities of survival over more than one interval (Table 11, column 7).

Conclusion

In table 12 is presented a summary of results from all four methods of estimation of cumulative risk.

Table 12. Summary of results of estimating cumulative risk over 5-year period using four different methods of estimation

Method	Direct 5-year cumulative risk	Indirect 5-year cumulative risk
Simple	0.8000	0.8125
Actuarial	0.8889	0.8429
Density	0.8731	0.8643
Kaplan Meier		0.8250

Since the most accurate measure is the Kaplan-Meier method we could compare all other results to this method. We could conclude that in this case study, the closest result to Kaplan-Meier method are obtained by indirect simple method, and by actuarial indirect method, while the most far away were results obtained by the direct actuarial method. One should be aware that this is not always the case. The results depend on the number of events and the number of censored cases. When the events are rare and there is no censoring, the discrepancy tends to be smaller (12).

Exercises

Data set

In Figure 3, another imaginary data-set is presented. Again, a cohort of 20 individuals initially without a disease under observation, were followed up for 5 years.

Task 1

For the data set presented in Figure 3, calculate cumulative risk using simple method:

- directly,
- indirectly by calculating annual risks first.

Task 2

For the data set presented in Figure 3, calculate cumulative risk using actuarial method:

- directly,
- indirectly by calculating annual risks first.

Task 3

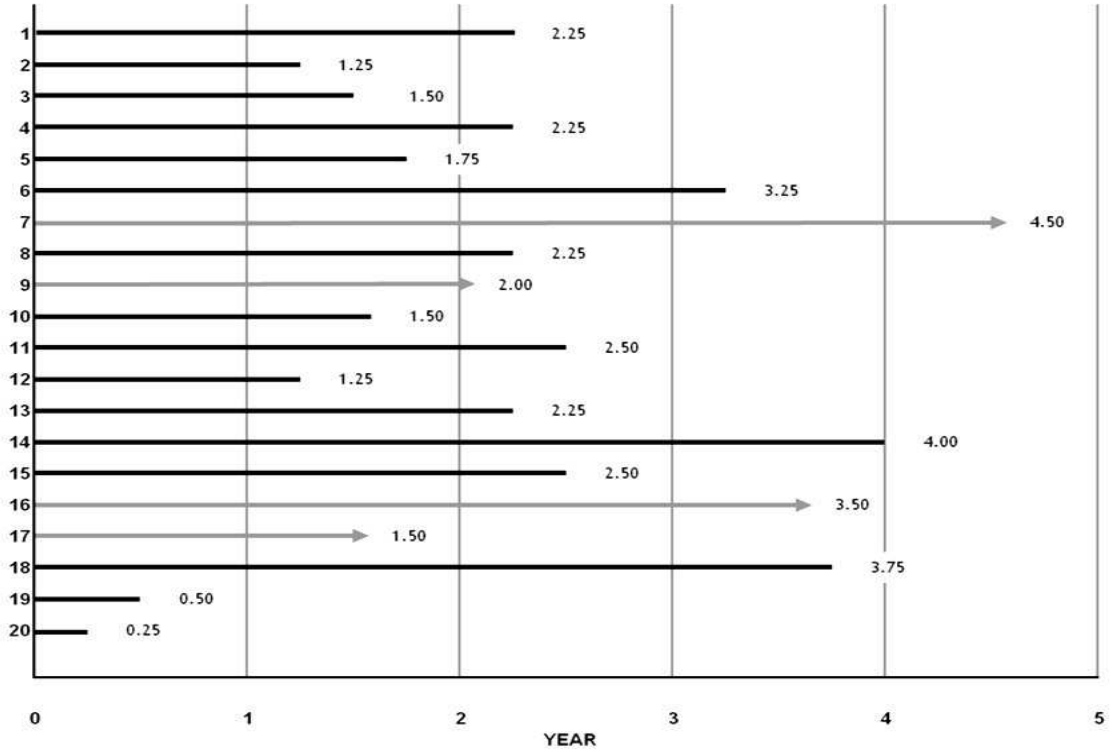
For the data set presented in Figure 3, calculate cumulative risk using density method:

- directly,
- indirectly by calculating annual risks first.

Task 4

For the data set presented in Figure 3, calculate cumulative risk using Kaplan-Meier method.

Figure 3. Graphic presentation of events in a cohort of 20 people



LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

References

1. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
2. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
3. Dos Santos Silva I. Cancer epidemiology: principles and methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 4, 2013.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
5. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Company; 1987.
6. Friis RH, Sellers TA. Epidemiology for public health practice. Fourth edition. Sudbury, MA: Jones&Bartlett Publishers; 2009.
7. Kue Young T. Population health: concepts and methods. Second edition. Oxford: Oxford University Press; 2005.
8. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
9. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
10. Kleinbaum DG, Klein M. Survival analysis. A Self-Learning Text. Second edition. New York: Springer; 2005.
11. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.

12. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
13. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 4, 2013.

Recommended readings

1. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
2. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.
3. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
4. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Measures of association and potential impact**
Module: 2.8	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Lijana Zaletel-Kragelj Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Measures of association, measures of potential impact.
Learning objectives	After completing this module students should: • know types of measures of association, and know how to calculate them; • know types of measures of potential impact, and know how to calculate them.
Abstract	Measures of association are quantities that express the strength of association between variables. Mathematically they are differences or ratios between different kinds of other measures (e.g. frequency measures). Measures of potential impact are quantities that express potential impact of one phenomenon on the frequency of the other. They are tightly interrelated. The module is describing basic features and types of them.
Teaching methods	An introductory lecture gives the students first insight in features and types of measures of association, and measures of potential impact. The theoretical knowledge is illustrated by two case studies. After introductory lectures students first carefully read the theoretical background of this module and complement their knowledge with recommended readings. Afterwards, on a provided data set, students perform in pairs two extensive tasks on calculation of different types of measures. Students are stimulated to compare results with their peers and discuss potential differences.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer for 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on calculated measures in which a detailed description of the process of calculation is described.

MEASURES OF ASSOCIATION AND POTENTIAL IMPACT

Lijana Zaletel-Kragelj

Theoretical background

Introduction to analysis of association and potential impact

Globally, measures of association, or measures of effect, are quantities that express the strength of association between variables (1-15). Mathematically, they are differences or ratios between different kinds of other measures (e.g. means, or frequency measures). Measures of potential impact are quantities that express potential impact of one phenomenon on the frequency of another (2,7,8,11). On one hand, they could express potential impact of risk factor on occurrence of observed health phenomena (unfavourable or favourable) in populations or among exposed persons. These measures are common in public health. On the other hand, they can express potential impact of an intervention on disease occurrence (benefit or harm) (16,17). These measures are much more commonly used in clinical epidemiology than in public health.

Before introducing some important measures of association and measures of potential impact, we need to provide, again, a warning about terminology. As we have already emphasized on several places, a common problem in epidemiology is the existence of multiple terms for the same concept. Also, there are instances where a single term is applied to different concepts. The confusion is aggravated by the multitude of terms that have been introduced, with usages that differ from one author to another (2).

Some important concepts in analysis of association

To make all considerations about analysis of associations between phenomena under observation easier, some concepts in relationship analysis should be clarified:

1. Observed phenomenon

   Observed phenomenon is a disease or other health-related condition under observation, frequently called also an outcome. In analysis of associations, it is assigned the role of "effect".
2. Risk factor

   Risk factor is defined as a phenomenon or characteristic (an aspect of behaviour or lifestyle, an environmental exposure, an inborn or inherited characteristic) which is, on the basis of epidemiologic evidence, known to be associated with health-related condition(s) considered important to prevent (1). Frequently it is called also an exposure. In analysis of associations, it is assigned the role of "cause".
3. Dependency

   In epidemiology, the dependent variable is the manifestation or outcome whose variation is to be explained by risk factors (1).

   As the relationship is not analyzed only by the means of epidemiologic methods but statistical methods as well, it should be clarified that in statistics the dependent variable is the one predicted by a regression equation.
4. Causality

   By definition, causality is relating causes to the effects they produce (1).

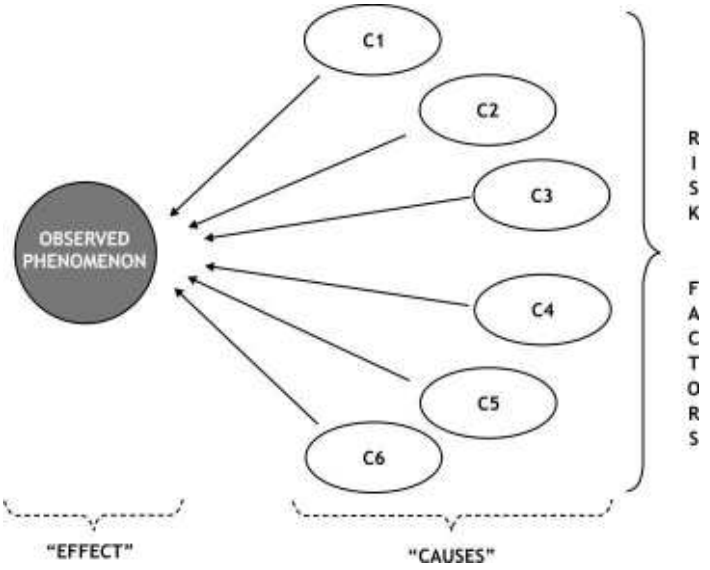
   Mostly, epidemiology is concerned with causality. In analysis of associations, the phenomenon under observation is usually assigned the role of "effect" and the risk factor thereof the role of "cause". Actually, causality of the relationships is not as simple. It should be clearly stated that epidemiologic evidence of causality by itself is insufficient to establish the causality in nature. The later is usually possible to be proved only by studies in different places which could be carried out for more decades. However, for the analysis of associations to be understandable, such setting up is very adequate. A picture (Figure 1) can offer additional help.

Some important considerations in analysis of association

Epidemiologic versus statistical measures

Distinguishing and separating between epidemiologic and statistical measures is rather difficult as they are interlaced, what is also discussed in one of chapters of this book. In fact, to some extent, statistical measures of association, used in epidemiology could be considered as epidemiologic and vice versa. Maybe the most appropriate attempt to distinguish between these two terms is as follows: when we are using the term "epidemiologic measure" this measure is applied to health phenomena of different kinds, while "statistical measure" is a more general term. "Epidemiologic measures" of association are also characterized by that they are based only upon frequency measures (e.g. relative risk, or odds ratio), while "statistical measures" are based upon frequency measures (e.g. chi-square test statistics), or mean and variance (e.g. test statistics in analysis of variance, regression and correlation coefficients). Only "epidemiologic measures" will be discussed here.

Figure 1. Graphical representation of the setting up of the observed disease or other health related phenomenon (observed phenomenon or "effect") and risk factors ("causes") into a relationship, as a help in analysis of associations



Simple versus complex analysis of association

In analyzing the relationship between observed outcome and risk factors two kinds of methods could be used:

1. Univariate methods

   Univariate methods of analysis of associations are concerned about the analysis of relationships of a single risk factor with a single outcome.

2. Complex or multivariate methods

   Multivariate methods are concerned about the analysis of relationships of several risk factors usually with a single outcome at a time. The methods for applying such kind of analysis are usually classified in statistical methods, although it is very difficult to draw a strict line of separation between statistical and epidemiologic methods. In fact, they all strive to assess the relationship between different phenomena.

   As outcomes are associated not only with one risk factor, it would be most appropriate to think about using multivariate methods as an analytical tool starting from the designing phase of a study. In fact, in epidemiology, multivariate methods are very useful in controlling the effect of confounders, in the phase of data analysis, since the methods of controlling them in the designing phase (randomization, matching) are not applicable in several types of epidemiologic studies (e.g. cross-sectional studies).

Interrelation between measures of association and measures of potential impact

Both, epidemiologic measures of association and measures of potential impact are based upon measures of frequency (e.g. incidence measures). They are tightly interrelated. As we will see later on, there is a case that the same measure is at the same time a measure of association, and a measure of potential impact. Only the interpretation is different. Additionally, measures of potential impact could be based upon measures of association as well.

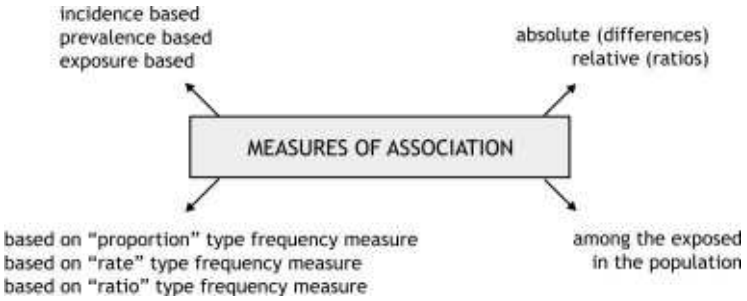
Likewise with measures of frequency, also these two families of measures could be classified according to various characteristics, what will be discussed later. The process of explanation of measures of both groups will be based upon example data in the case studies after some theoretical background.

Measures of association

Epidemiologic measures of association are quantities that express the strength of association between phenomena, related to health (1-15). Mathematically, they are differences or ratios between different kinds of frequency measures.

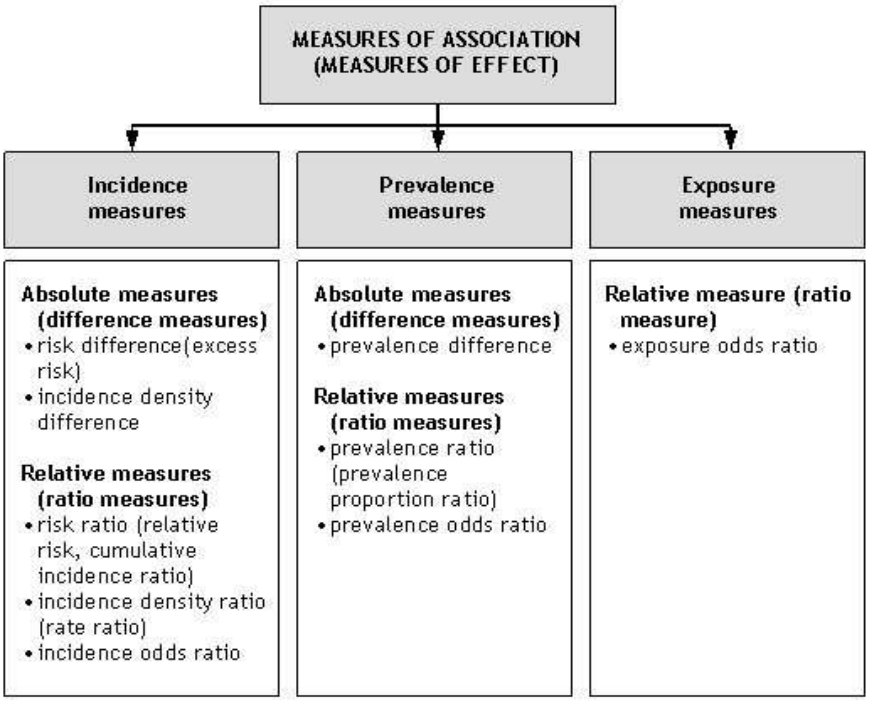
Measures of association could be classified according to various characteristics. Four different classifications are summarized in Figure 2.

Figure 2. Classifications of measures of association according to various characteristics



1. Classification of measures of association to incidence-based, prevalence-based, and exposure-based
   According to this, if the measure is incidence or prevalence-based, the measures of association are classified as incidence-based and prevalence-based. All measures have been already introduced, so in this place they are only listed by this criterion.
2. Classification of measures of association to absolute and relative
   Absolute measures are differences between frequency measures of different kinds between two observed groups (e.g. between exposed and unexposed), while relative measures are ratios between frequency measures of different kinds between two observed groups (e.g. between exposed and unexposed).
3. Classification of measures of association according to the type of relative frequency measure upon which are based
   Measures of association could be based on different types of frequency measures that were in details presented in one of previous modules in this book. According to this classification, they could be based on proportion (measure in which numerator is included in the denominator in basic frequency measure), rate or ratio in a narrow sense frequency measures (measures in which numerator is not included in the denominator in basic frequency measure) (4).

Figure 3. Measures of association frequently used in epidemiologic research



4. Classification of absolute measures of association to those related to exposed and those related to population
   Absolute measures of association could be classified to those related to exposed, and those related to population. This classification is not so important for measures of association compared with the measures of potential impact. Consecutively, we will not pay attention to this classification at the moment.

Detailed interrelation between different measures of association are out of the scope of this module and are discussed in several textbooks of modern epidemiology (2,4,8,18). In continuation, the measures of association frequently used in epidemiologic research (Figure 3) will be presented.

Incidence comparisons

In incidence (cohort) studies, incidence measures are generated, being absolute (difference-based) or relative (ratio-based).

Difference-based measures

Two difference-based measures are commonly used in incidence comparisons: risk difference and incidence density difference.

1. **Risk difference or excess risk**

   This measure is the absolute difference between two risks (1,5,12). When the risk difference is observed between exposed and unexposed (excess risk among exposed), the measure is calculated as follows (1,5,12) (Equation 1):

$$RD = R_{E+} - R_{E-} \qquad \textbf{Equation 1.}$$

   RD = risk difference among exposed
   R_{E+} = risk among exposed
   R_{E-} = risk among unexposed

   The risk difference can be estimated directly from the cumulative incidence difference (2,4) (Equation 2):

$$CID = CI_{E+} - CI_{E-} \qquad \textbf{Equation 2.}$$

   CID = cumulative incidence difference among exposed
   CI_{E+} = cumulative incidence among exposed
   CI_{E-} = cumulative incidence among unexposed

   Calculation of this measure in practice is presented in Case study 1.

2. **Incidence density difference** (2,4,12)

   It is the absolute difference between two incidence densities (1). The measure is calculated as follows (Equation 3):

$$IDD = ID_{E+} - ID_{E-} \qquad \textbf{Equation 3.}$$

   IDD = incidence density difference
   ID_{E+} = incidence density among exposed
   ID_{E-} = incidence density among unexposed

   Calculation of this measure in practice is presented in Case study 1.

Ratio-based measures

Following, the ratio-based measures are commonly used in incidence comparisons: relative risk or risk ratio, incidence density ratio, and risk odds ratio.

1. **Risk ratio or relative risk**

   Relative risk is a ratio of the risk of observed outcome (e.g. disease, death) among the exposed to the risk among the unexposed (2,4,5,7,8,12). The measure is calculated as follows (Equation 4):

$$RR = \frac{R_{E+}}{R_{E-}} \qquad \textbf{Equation 4.}$$

   RR = relative risk
   R_{E+} = risk among exposed
   R_{E-} = risk among unexposed

   Calculation of this measure is presented in Case study 1.

2. **Incidence density ratio**

   Incidence density ratio or rate ratio is the ratio of two incidence densities (2,4,7,12). The measure is calculated as follows (Equation 5):

$$IDR = \frac{ID_{E+}}{ID_{E-}} \qquad \textbf{Equation 5.}$$

   IDR = incidence density ratio
   ID_{E+} = incidence density among exposed
   ID_{E-} = incidence density among unexposed

   Calculation of this measure is presented in Case study 1.

Herewith, we need to introduce a new ratio effect measure – a measure known under the term "hazard ratio" (2,19,20). This measure is tightly related to the incidence density ratio (19). It is a very important measure since it is the measure of effect in survival analysis (19,20), which is closely related to epidemiologic concept of risk, what is presented in detail in a separate module in this book. To understand the relationship between incidence density ratio and hazard ratio, we need to recall the measure known as "hazard rate". This index, that measures the instantaneous potential for change in disease status (4,20,21), has been presented in detail in a separate module in this book. Incidence density is an average rate for estimating average of instantaneous incidence rates. Under certain conditions we could use the terms "incidence density", and "hazard rate" as synonyms (19). The incidence density ratio, just presented, is the ratio of two incidence densities (two incidence rates). Conceptually, this ratio is identical to a hazard ratio usually denoted as HR (19). Theoretically, the hazard ratio at a given point in time is the limiting value of the incidence density ratio as the time around the point becomes very short, just as the hazard is the limiting quantity for incidence density (19).

3. **Incidence odds ratio**

Incidence or risk odds ratio is a ratio of risk odds of observed outcome (e.g. disease, death) among the exposed to the odds among the unexposed (2,4,5,7,8,12). The measure is calculated as follows (Equation 6). Calculation of this measure is presented in Case study 1.

Odds ratio is one of extremely useful measures in multivariate analysis as well. It is one of possible results of logistic regression methods (22,23).

$$ROR = \frac{{}_{risk}O_{E+}}{{}_{risk}O_{E-}} \qquad \textbf{Equation 6.}$$

ROR= risk odds ratio

${}_{risk}O_{E+}$ = risk odds among exposed

${}_{risk}O_{E-}$ = risk odds among unexposed

Herewith, we need to expose the relationship between relative risk and odds ratio. In the medical literature, odds ratio is often misinterpreted as estimate of relative risk. Odds ratio is a good estimate of relative risk only under specific conditions - only when the phenomenon under observation is rare. In this case only, denominators in the calculation process of frequency measures for the odds ratios (e.g. number of people under observation without observed phenomenon among exposed) and for the risk ratios (e.g. total number of exposed), are similar. Consequently, values of both measures, the odds ratio and relative risk, are very close to each-other.

The ratio measures values which range from 0 to infinity. Values close to 1.0 indicate no relationship between the exposure and the outcome. Values less than 1.0 suggest a protective effect, while values greater than 1.0 suggest an adverse effect of exposure. When comparing all three ratio measures it can be shown that, numerically, the odds ratio falls the furthest from the null, and the risk ratio the closest, with the rate ratio (incidence density ratio) in-between (19).

Prevalence comparisons

In prevalence (cross-sectional) studies, prevalence measures are generated, being, similarly as in incidence studies, absolute (difference-based), or relative (ratio-based).

Difference-based measures

Only one difference-based measure is commonly used in prevalence comparisons.

1. **Prevalence proportion difference**

Prevalence proportion difference, also known as prevalence rate difference (2,4,19), is calculated as follows (4) (Equation 7):

$$PD = P_{E+} - P_{E-} \qquad \textbf{Equation 7.}$$

PD = prevalence difference among exposed

P_{E+} = prevalence proportion among exposed

P_{E-} = prevalence proportion among unexposed

Calculation of this measure is presented in Case study 1.

Ratio-based measures

Following, two ratio-based measures are commonly used in prevalence comparisons: prevalence ratio, and prevalence odds ratio.

1. **Prevalence ratio**

Prevalence ratio is a ratio of point prevalence proportion of observed outcome among the exposed to the point prevalence proportion among the unexposed (4,8,12,15). The measure is calculated as follows (Equation 8):

$$PR = \frac{P_{E+}}{P_{E-}} \qquad \textbf{Equation 8.}$$

PR = prevalence ratio
P_{E+} = prevalence (proportion) among exposed
P_{E-} = prevalence (proportion) among unexposed

Calculation of this measure is presented in Case study 1.

2. **Prevalence odds ratio**
Prevalence odds ratio is a ratio of prevalence odds of observed outcome among the exposed to the prevalence odds among the unexposed (2,4,8,12,15). The measure is calculated as follows (Equation 9):

$$POR = \frac{{}_{prevalence}O_{E+}}{{}_{prevalence}O_{E-}} \qquad \textbf{Equation 9.}$$

POR= prevalence odds ratio
${}_{prevalence}O_{E+}$ = prevalence odds among exposed
${}_{prevalence}O_{E-}$ = prevalence odds among unexposed

Calculation of this measure is presented in Case study 1.

Exposure comparisons

In case-control studies neither incidence nor prevalence measures can be generated. The only measures that can be calculated are exposure-related measures. The exposure odds ratio is the ratio of the odds in favour of exposure among the cases to the odds in favour of exposure among non-cases (1,2,4,8). The measure is calculated as follows (Equation 10):

$$EOR = \frac{{}_{exposure}O_{cases}}{{}_{exposure}O_{non-cases}} \qquad \textbf{Equation 10.}$$

EOR= exposure odds ratio
${}_{exposure}O_{cases}$ = exposure odds among cases
${}_{exposure}O_{non\text{-}cases}$= exposure odds among non-cases

Calculation of this measure is presented in Case study 1.

Major advance is that exposure odds ratio (EOR) and disease odds ratio (DOR) in case-control studies are mathematically equivalent. Consecutively, the exposure odds ratio can be used to estimate the relative risk, especially when the probability of positive response is small (2,4,8). The equivalence is presented in Case study 1.

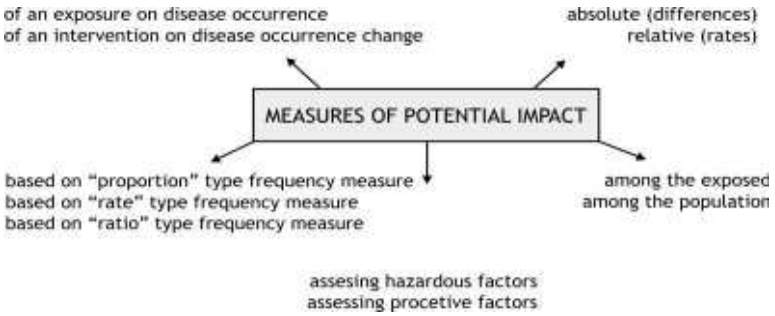
Measures of potential impact

This family of epidemiologic measures quantifies the potential impact of various exposures on observed phenomena. They could express the potential impact of risk factors on occurrence of observed health phenomena or the potential impact of an intervention on disease occurrence change (1,2,4,5,13,16,24,25)

Measures of potential impact could be classified according to various characteristics. Five different classifications are summarized in Figure 4.

1. Classification of measures of potential impact to those assessing impact of an exposure on disease occurrence, and those assessing impact of an intervention on disease occurrence change
In this group of measures, there are two subgroups according to what they are measuring. Measures of the first subgroup express the potential impact of risk factor on occurrence of observed health phenomena among exposed persons, or in the population. These measures are common in public health. Measures of the second subgroup express the potential impact of an intervention on disease occurrence reduction. They are much more common in clinical epidemiology than in public health.
2. Classification of measures of potential impact to absolute and relative
According to this, if the measure is expressed as difference or ratio, the measures of potential impact are classified as absolute (differences) and relative (ratios).

Figure 4. Classifications of measures of potential impact according to various characteristics



3. Classification of measures of potential impact according to measures referring to the population or to exposed

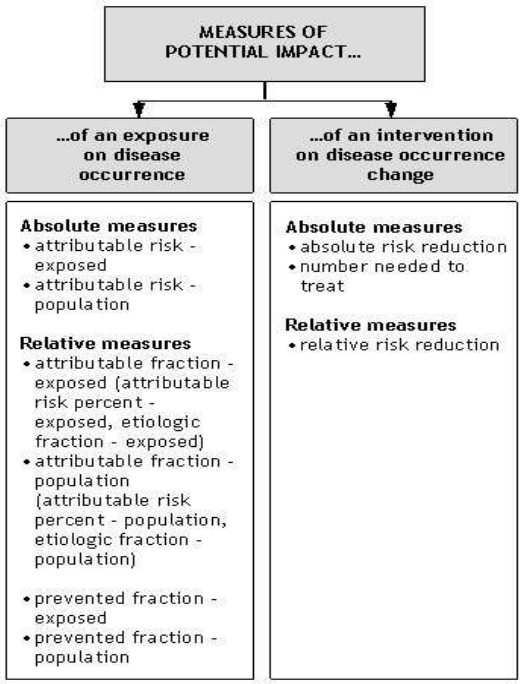
   Measures of potential impact could be classified also according to those related to exposed, and those related to population.
4. Classification of measures of potential impact according to the type of relative frequency measure upon which they are based

   Measures of potential impact could be based on different types of frequency measures that were presented in detail in one of previous modules in this book. According to this classification, they can be based on proportion (measure in which numerator is included in the denominator in basic frequency measure), rate or ratio in a narrow sense frequency measures (measures in which numerator is not included in the denominator in basic frequency measure) (4).
5. Classification of measures of potential impact according to measures assessing hazardous and measures assessing protective factors

   Measures dealing with hazardous factors are known under the common term "etiologic fraction" (4). The other group of measures, those measures that deal with protective risk factors, are known under the common term "preventable fraction" (4). The difference is that they express the proportion of potential new cases of the disease under observation that would have occurred in absence of exposure to a protective factor but did not occur (4), or the proportion of the hypothetical total load of disease that has been prevented by exposure to a protective factor (1).

Detailed interrelation between different measures of potential impact are out of the scope of this module and are discussed in several textbooks of modern epidemiology (2,4,8,18). In continuation, there will be presented measures of association frequently used in epidemiologic research (Figure 5).

Figure 5. Measures of potential impact frequently used in epidemiologic research



Measures assessing impact of an exposure on disease occurrence

Difference-based measures

Difference-based measures of potential impact are: attributable risk in exposed and attributable risk in population.

1. **Attributable risk in exposed**

   This measure is the absolute difference between two risks (1,2,5,8). The measure is calculated as follows (Equation 11):

$$AR = R_{E+} - R_{E-} \qquad \textbf{Equation 11.}$$

   AR = attributable risk among exposed
   R_{E+} = risk among exposed
   R_{E-} = risk among unexposed

In fact, this measure has been already introduced among measures of association under the name "risk difference". It is the portion of the risk in the exposed that is due to the exposure (could be attributed to exposure). In other terms, it is the risk of a disease in the exposed that could be eliminated if exposure were eliminated.

Calculation of this measure is presented in Case study 2 (technically, it is exactly the same as calculation of risk difference presented in Case study 1, but the interpretation is different).

In the calculation process, it is assumed that risk factors other than the one under observation have equal effects on the exposed and unexposed.

2. **Attributable risk in the population**

Attributable risk could be expressed also for the population (population attributable risk). It is a measure of the amount of disease attributed to a putative cause of the disease in the population (1,2,5,8). Mathematically, it is the difference between the risk of disease in the entire population and among the unexposed (Equation 12):

$$PAR = R_{pop} - R_{E-} \qquad \textbf{Equation 12.}$$

PAR = population attributable risk
R_{pop} = risk in population
R_{E-} = risk among unexposed

Population attributable risk is the portion of the risk of a disease in the population (exposed and non-exposed) that is due to exposure. It is the risk of a disease in the population that could be eliminated if exposure were eliminated. Calculation of this measure is presented in Case study 2.

Ratio-based measures – assessment of hazardous factors

Ratio-based measures of potential impact of hazardous factors are the attributable fraction in the exposed and attributable fraction in the population.

1. **Attributable fraction in the exposed**

It is a fraction of people with the disease under observation which could be attributed to exposure to a risk factor under observation (1,2,5,7,8). It is known also under the terms attributable risk proportion (2), and etiologic fraction (4). The measure is calculated as follows (Equation 13):

$$AF = \frac{R_{E+} - R_{E-}}{R_{E+}} \qquad \textbf{Equation 13.}$$

AF = attributable fraction among exposed
R_{E+} = risk among exposed
R_{E-} = risk among unexposed

In the calculation process, it is assumed that risk factors other than the one under observation have equal effects on the exposed and unexposed.

When this measure is multiplied by a multiplier 100 it is known as attributable risk percent (5).

Attributable fraction or attributable risk percent is the fraction/percent of the risk of a disease in the exposed that is due to the exposure. It is the fraction/percent of the risk of a disease in the exposed that could be eliminated if exposure were eliminated.

The same measure could be calculated using another procedure (1) (Equation 14):

$$AF = \frac{RR - 1}{RR} \qquad \textbf{Equation 14.}$$

AF = attributable fraction among exposed
RR = relative risk

Calculation of this measure is presented in Case study 2, as well.

2. **Attributable fraction in the population**

When attributable fraction refers to the population (attributable fraction – population, or population attributable fraction), the measure is calculated as follows (1) (Equation 15):

$$PAF = \frac{R_{pop} - R_{E-}}{R_{pop}} \qquad \textbf{Equation 15.}$$

PAF = attributable fraction in population
R_{pop} = risk in population
R_{E-} = risk among unexposed

When this measure is multiplied by a multiplier 100 it is known as population attributable risk percent (5).

Population attributable fraction or percent is the fraction/percent of the risk of a disease in the population (exposed and non-exposed) that is due to exposure. It is the fraction/percent of the risk of a disease in the population that would be eliminated if exposure were eliminated.

The same measure could be calculated using another procedure (Equation 16). Calculation of this measure is presented in Case study 2.

$$PAF = \frac{P_{E+} \times (RR - 1)}{1 + [P_{E+} \times (RR - 1)]}$$ **Equation 16.**

PAF = attributable fraction in population
RR = relative risk
P_{E+} = proportion of exposed in population

Ratio-based measures – assessment of protective factors

When considering a protective exposure, appropriate alternative measures are prevented fraction in the exposed and prevented fraction in the population (1,19,26).

1. **Prevented fraction in the exposed**

The reference point in this measure is the risk of disease if nobody is exposed (26). The fraction indicates the amount of disease that would be prevented by exposure to a protective factor (26). The measure is calculated as follows (Equation 17). In calculation process it is assumed that risk factors other than the one under observation have equal effects on the exposed and unexposed. Calculation of this measure is presented in Case study 2. However, the same measure could be calculated using another procedure (26) (Equation 18). Calculation using this procedure is presented in Case study 2, as well.

$$PF = \frac{R_{E-} - R_{E+}}{R_{E-}}$$ **Equation 17.**

PF = prevented fraction among exposed
R_{E-} = risk among unexposed
R_{E+} = risk among exposed

$$PF = 1 - RR$$ **Equation 18.**

PF = prevented fraction among exposed
RR = relative risk

This measure is identical to the measure, known in clinical epidemiology as "relative risk reduction" (26). It will be discussed later in this module. It is also identical to the measure known in communicable diseases epidemiology as "vaccine efficacy" (26).

2. **Prevented fraction in the population**

When the prevented fraction refers to the population (prevented fraction - population), the measure is calculated as follows (26) (Equation 19):

$$PPF = \frac{R_{E-} - R_{pop}}{R_{E-}}$$ **Equation 19.**

PPF = prevented fraction in population
R_{E-} = risk among unexposed
R_{pop} = risk in population

Calculation of this measure is presented in Case study 2.

The same measure could be calculated using another procedure (26) (Equation 20). Calculation using this procedure is presented in Case study, 2 as well.

$$PPF = P_{E+} \times (1 - RR)$$ **Equation 20.**

PPF = prevented fraction in population
RR = relative risk
P_{E+} = proportion of exposed

Beside these two measures, there exists also a measure called "preventable fraction". This measure is less commonly used and will not be discussed here.

Measures assessing impact of an intervention on disease occurrence change

As already mentioned, this group of measures is much more commonly used in clinical epidemiology than in public health since they measure the effectiveness of clinical intervention trials. Nevertheless, they could be effectively used in assessment of public health interventions, as well.

There exist several subgroups of these measures, depending on if they are assessing undesirable ("bad" outcomes) or desirable events ("good" outcomes) (16). In this module we are presenting only three measures used most frequently: absolute risk reduction, relative risk reduction, and number needed to treat.

1. **Absolute risk reduction**

Absolute risk reduction is defined as the difference in risk between the control (untreated) group and the intervention (treated) group (25). The measure is calculated as follows (Equation 21):

$$ARR = R_C - R_T \qquad \textbf{Equation 21.}$$

ARR = absolute risk reduction
R_C = risk in the control group
R_T = risk in the intervention (treated) group

This measure expresses the proportion of individuals spared from the unfavourable outcome if they receive the intervention in comparison to not receiving it (25). In fact, this measure is the opposite of the attributable risk in the exposed (Equation 11). This calculation is opposite since clinical interventions on general reduce risk (26). Calculation of this measure is presented in Case study 2.
Herewith, we need to emphasize that in clinical epidemiology, another notation of basic frequency measures is frequently used. Risk in the control group is denoted by CER (the control group event rate), while risk in the intervention (treated) group is denoted by EER (experimental group event rate). This notation is suitable since the studies conducted in clinical epidemiology are mostly clinical trials (experimental studies with the experimental and the control group).

2. Relative risk reduction

This measure is the most commonly reported measure in this family of measures (25). It can be obtained by dividing the absolute risk reduction by the risk in the control group (25) (Equation 22):

$$RRR = \frac{R_C - R_T}{R_C} \qquad \textbf{Equation 22.}$$

RRR = relative risk reduction
R_C = risk in the control group
R_T = risk in the intervention (treated) group

In the calculation process it is assumed that risk factors other than the one under observation have equal effects on the exposed and the unexposed.

Relative risk reduction measures how much of the risk is reduced in the experimental (treated) group compared to a control group. Here would be worthy to mention, that treatments with very large relative risk reductions may have a small effect in conditions where the control group has a very low negative outcome. On the other hand, a modest relative risk reduction can mean major clinical importance if the risk in a control group is large. Calculation of this measure is presented in Case study 2.

$$RRR = 1 - RR = 1 - \frac{R_T}{R_C} \qquad \textbf{Equation 23.}$$

RRR = relative risk reduction
RR = relative risk
R_C = risk in the control group
R_T = risk in the intervention (treated) group

The same measure could be obtained easily from the relative risk (the ratio of risk in the intervention group to the risk in the control group) using the equation 23 (25). Calculation using this procedure is presented in Case study 2, as well. This measure expresses how much the risk is reduced in the treated group compared to a control group. The greater the relative risk reduction, the more efficacious is the intervention (25).

3. Number needed to treat

The last measure in this group that we present here is the number needed to treat. It is in fact only another way to express the absolute risk reduction, since it is defined as the inverse of the absolute risk reduction (25) (Equation 24). Calculation of this measure is presented in Case study 2.

$$NNT = \frac{1}{ARR} = \frac{1}{R_C - R_T} \qquad \textbf{Equation 24.}$$

NNT = number needed to treat
ARR = absolute risk reduction
R_C = risk in the control group
R_T = risk in the intervention (treated) group

This measure is very popular because of its simplicity to compute and its ease to interpret – it is interpreted as the number of patients that would need to be treated to prevent an additional negative outcome.

Case studies

Case study 1: Measures of association

Incidence comparisons

An incidence study of impact of hyperirritable uterus on a premature delivery

In an ambidirectional cohort study, based on Perinatal Informational System of Slovenia (PISS)[6] (27), data of 800 mothers and their newborns were analyzed.

The observed outcome was a preterm delivery (a delivery prior the completed 36th week of pregnancy). We related this outcome to a condition ("exposure"), known as hyperirritable uterus (abnormal contractility of the uterus during earlier stages of pregnancy). For analyzing this relationship we used different methods:

1. First, we were interested only if preterm delivery occurred or not. The results of this simple cumulative method analysis are presented in Table 1.

Table 1. Frequency of preterm delivery in two groups of mothers according to absence or presence of hyperirritable uterus during pregnancy based on Slovene PISS data (27)

	Presence of hyperirritable uterus during pregnancy		Total
Preterm delivery	**No**	**Yes**	
No	719	30	749
Yes	44	7	51
Total	763	37	800

Table 2. The intermediate measures (frequency measures) - necessary elements for calculation of incidence risk and odds-based measures of association calculated on the basis of data from Table 1

Frequency measure	Notation	Calculation and value
Risk in the non-exposed	R_{E-}	$R_{E-} = \frac{44}{763} = 0.0577$
Risk in the exposed	R_{E+}	$R_{E+} = \frac{7}{37} = 0.1892$
Risk (incidence) odds in the non-exposed	${}_{risk}O_{E-}$	${}_{risk}O_{E-} = \frac{44}{719} = 0.0612$
Risk (incidence) odds in the exposed	${}_{risk}O_{E+}$	${}_{risk}O_{E+} = \frac{7}{30} = 0.2333$

The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 1 – are presented in Table 2.

2. Afterwards, we were also interested in the time when preterm delivery occurred. We used the data on duration of pregnancy (assessed through gestational age of a newborn). These more detailed data are presented in Table 3. The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 3 – are presented in Table 4.

[6] In Slovenia, already for several years, for the purpose of teaching epidemiologic methods in public health, comprising also statistical methods, data collection which enables learning such methods in much comprehensive way has been created. These are the data collected in the frame of the Perinatal Informational System of Slovenia (PISS) (27), which is considered to be one of the permanent health data-bases of the highest quality with many years' tradition in the country. It was initiated in 1987, when collection of perinatal data started according to a common protocol in all fourteen Slovene maternity hospitals.

In PISS, the data on total period of pregnancy and on delivery are joined. Data concerning the pregnancy period are collected at the time of admission to the maternity. They are objective and subjective. Objective data come from a so-called "Maternal Booklet". This is a booklet every pregnant woman in Slovenia gets at the time of established pregnancy. It holds the most important up-to-date information regarding pregnancy, determined by the obstetrician, responsible for a pregnant woman. This information is important in case of an urgent treatment or intervention and at the time of delivery. Subjective data come from a questionnaire filled-in at the time of admission to the maternity. They are collected by the maternity staff (partially by a midwife and partially by an obstetrician). After admission, a period from the admission to the discharge from a maternity is prospectively followed-up and data collected in PISS.

The basic data material for all epidemiologic and statistical activities is composed of 6,356 statistical units, representing the model of a population. For teaching different epidemiologic and statistical methods, samples of various sizes are then randomly selected from the population database. The data set used in this example is composed of 800 observational units.

Data material for teaching is only a small piece out of the whole collection PISS, prepared especially for this purpose. Safeguard of personal data is assured so that all personal data have been removed, and moreover, the data are selected from the whole collection which shall be used only for the teaching purpose.

Calculation of measures of association in incidence comparisons

The procedures of calculation of risk difference, incidence density difference, relative risk, incidence density ratio (also known as rate ratio), and incidence odds ratio (also known as risk odds ratio), and interpretation of results is as follows:

1. Risk difference or excess risk

The risk difference calculated according to Equation 1 based on data presented in Tables 1 and 2 is as follows (Equation 25):

$$RD = 0.1892 - 0.0577 = 0.1315 \qquad \textbf{Equation 25.}$$

Mothers with hyperirritability of uterus have a risk of preterm delivery 0.1315 higher than mothers without it.

We could interpret this result also in terms of a percentage. If we multiply the risk difference by 100, we get 13.15%, and the interpretation is that mothers with hyperirritability of uterus have a risk of preterm delivery 13.1% higher than mothers without it.

2. Incidence density difference

The incidence density difference calculated according to Equation 3 based on data presented in Tables 3 and 4 is as follows (Equation 26):

$$IDD = 0.00503 - 0.00147 = 0.00355 \qquad \textbf{Equation 26.}$$

Mothers with hyperirritability of uterus have an incidence density of preterm delivery 0.00355 higher than mothers without it.

If we than multiply the incidence density difference by 1,000 we get the value 3.55, which could be interpreted as follows: on average, in 42-week interval, 3.55 more pregnant women with hyperirritability of the uterus experience preterm delivery per 1,000 population than mothers without it.

3. Risk ratio or relative risk

The relative risk calculated according to Equation 4 based on data presented in Tables 1 and 2 is as follows (Equation 27):

$$RR = \frac{0.1892}{0.0577} = 3.2807 \qquad \textbf{Equation 27.}$$

The risk of preterm delivery in mothers with hyperirritability of uterus is 3.28-times higher than in mothers without it.

Table 3. Duration of pregnancy and person-weeks quantity in two groups of mothers according to absence or presence of hyperirritable uterus during pregnancy based on Slovene PISS data (27)

Hyperirritable uterus during pregnancy no			Hyperirritable uterus during pregnancy yes		
Duration of pregnancy (weeks of gestation)	**n**	**Person-weeks (PW)**	**Duration of pregnancy (weeks of gestation)**	**n**	**Person-weeks (PW)**
22	1	22	28	1	28
23	1	23	29	2	58
25	1	25	32	1	32
26	1	26	33	1	33
29	2	58	34	1	34
30	2	60	35	1	35
31	2	62	37	5	185
32	7	224	38	3	114
33	2	66	39	9	351
34	5	170	40	11	440
35	12	420	41	2	82
36	8	288	Total	37	1392
37	33	1221			
38	97	3686			
39	182	7098			
40	295	11800			
41	108	4428			
42	4	168			
Total	763	29845			

Table 4. The intermediate measures (frequency measures) - necessary elements for calculation of incidence density based measures of association calculated on the basis of data presented in Tables 1 and 3

Frequency measure	Notation	Calculation and value
Incidence density in the non-exposed	ID_{E-}	$ID_{E-} = \frac{44}{29{,}845} = 0.00147$
Incidence density in the exposed	ID_{E+}	$ID_{E+} = \frac{7}{1{,}392} = 0.00503$

4. Incidence density ratio

The incidence density ratio calculated according to Equation 5 based on data presented in Tables 3 and 4 is as follows (Equation 28):

$$IDR = \frac{0.00503}{0.00147} = 3.4110 \qquad \textbf{Equation 28.}$$

The "risk" of preterm delivery in mothers with hyperirritability of uterus, assessed through incidence density procedure, is 3.41-times higher than in mothers without it.

5. Incidence or risk odds ratio

The risk odds ratio calculated according to Equation 6 based on data presented in Tables 1 and 2 is as follows (Equation 29):

$$ROR = \frac{0.2333}{0.0612} = 3.8129 \qquad \textbf{Equation 29.}$$

The odds of preterm delivery in mothers with hyperirritability of uterus is 3.81-times higher than in mothers without it.

When comparing all the three presented ratio measures for the same set of data, we can see that the odds ratio falls the furthest from the null (ROR = 3.81), and the relative risk the closest (RR = 3.28), with the incidence density ratio in-between them (IDR = 3.41).

Prevalence comparisons

A cross-sectional study of smoking in adults in Slovenia

In a cross-sectional study, basing on CINDI Health Monitor survey 2001, that was aiming at assessing the prevalence of health behaviours (28)[7], data of 9,034 adults were analyzed.

The observed outcome was smoking at the time (point of observation) of the survey. We related this outcome to a gender as a risk factor for unfavourable health behaviour (i.e. in this context in a role of "exposure"; since, in general, males are at the highest risk for smoking than females, they are considered as "exposed", while females are considered as "non-exposed").

Out of 9,034 respondents, 8,904 adults reported on their smoking status. This outcome was related to gender. The results of this analysis are presented in Table 5.

The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 5 – are presented in Table 6.

[7] In Slovenia in 2001 the survey aiming at assessing the prevalence of health behaviours (stress perception, smoking habits, nutrition habits, alcohol consumption habits, physical activity habits, and traffic safety habits) was performed for the first time. This survey is conceptually a part of a wider international project in the frame of the Countrywide Integrated Non-communicable Diseases Intervention (CINDI) program, entitled CINDI Health Monitor Surveys, supported by the World Health Organization, CINDI Health Monitor. For the second and the third time this type of a survey was performed in 2004 and in 2008.

In the 2001 survey, the stratified random sample was drawn from the Central Population Registry of the Republic of Slovenia. The sample size was 15,379 with the age range 25-64 years. The sampling was performed by the Statistical Office of the Republic of Slovenia.

The data were collected in late spring 2001 by using a self-administered postal questionnaire, based on the CHM Core Questionnaire (28).

Out of 15,379 inhabitants included in the sample 15,153 were contacted (226 were excluded because of changed domicile, severe illness or death). The response rate was 63.8% (9,666 responses). The respondents did not differ statistically from non-respondents in age distribution or distribution of size of settlements of permanent residence, but the response to the survey was slightly lower among men (47.0%) than among women (53.0%) at a ratio 1:1.1 (according to population data in 2001 the ratio was 1:1). The questionnaires of 9,034 respondents were eligible for analysis (eligibility criteria: sex and age provided by Statistical Office of the Republic of Slovenia).

For the purpose of this module, we have chosen observation of smoking behaviour.

Table 5. Prevalence of smoking in both gender groups in adult population in Slovenia in 2001, based on CINDI Health Monitor Survey, Slovenia 2001 (28)

Smoking	Gender		Total
	Females	Males	
No	3,859	2,931	6,790
Yes	971	1,143	2,114
Total	4,830	4,074	8,904

Table 6. The intermediate measures (frequency measures) - necessary elements for calculation of prevalence measures of association calculated on the basis of data presented in Table 5

Frequency measure	Notation	Calculation and value
Prevalence proportion in the non-exposed (females)	P_{E-}	$P_{E-} = \frac{971}{4,830} = 0.2010$
Prevalence proportion in the exposed (males)	P_{E+}	$P_{E+} = \frac{1,143}{4,074} = 0.2806$
Prevalence odds in the non-exposed (females)	${}_{prevalence}O_{E-}$	${}_{prevalence}O_{E-} = \frac{971}{3,859} = 0.3900$
Prevalence odds in the exposed (males)	${}_{prevalence}O_{E+}$	${}_{prevalence}O_{E+} = \frac{1,143}{2,931} = 0.2333$

Calculation of measures of association in prevalence comparisons

The procedures of calculation of prevalence proportion difference, prevalence ratio (also known as prevalence rate ratio), and prevalence odds ratio, and interpretation of results is as follows:

1. Prevalence proportion difference

   The prevalence proportion difference calculated according to Equation 7 based on data presented in Tables 5 and 6 is (Equation 30):

   $$PD = 0.2806 - 0.2010 = 0.0796$$ **Equation 30.**

   In Slovene adult males, prevalence of smoking is 0.0796 higher than in Slovene adult females.

   We could interpret this result also in terms of a percentage. If we multiply the prevalence proportion difference by 100, we get 7.96%, and the interpretation is that, in Slovene adult males the prevalence of smoking is 7.96% higher than in Slovene adult females.

2. Prevalence ratio

   The prevalence ratio calculated according to Equation 8 based on data presented in Tables 5 and 6 is as follows (Equation 31):

   $$PR = \frac{0.2806}{0.2010} = 1.3956$$ **Equation 31.**

   The prevalence of smoking in Slovene adult males is 1.40 times higher than in Slovene adult females.

3. Prevalence odds ratio

   The prevalence odds ratio calculated according to Equation 9 based on data presented in Tables 5 and 6 is as follows (Equation 32):

   $$POR = \frac{0.3900}{0.2516} = 1.5498$$ **Equation 32.**

The (prevalence) odds of smoking in Slovene adult males, is 1.55 times higher than in Slovene adult females.

Exposure comparisons

A case-control study of impact of hyperirritable uterus on premature delivery

On the basis of the Slovene PISS data (27), we could simulate a case-control study. Let's suppose that we have followed-up outcomes of deliveries in a given period of time in a selected maternity. In this period we registered 51 preterm deliveries. For each case we selected two controls out of mothers without preterm delivery (i.e. on-time delivery). We were interested whether the frequency of hyperirritability of

uterus (exposure) was different in cases (mothers with preterm delivery) in comparison to controls (mothers without preterm delivery). The results of this analysis are presented in Table 7.

Table 7. Frequency of preterm delivery in two groups of mothers according to absence (controls) or presence (cases) of hyperirritable uterus during pregnancy based on Slovene PISS data (27)

	Presence of hyperirritable uterus during pregnancy		Total
Preterm delivery	**No (Non-exposed)**	**Yes (Exposed)**	
No (Controls)	95	7	102
Yes (Cases)	44	7	51
Total	139	14	153

Table 8. The intermediate measures (frequency measures) - necessary elements for calculation of measures of association in exposure comparisons calculated on the basis of data presented in Table 7

Frequency measure	Notation	Calculation and value
Exposure odds in controls	${}_{exposure}O_{controls}$	${}_{exposure}O_{controls} = \frac{7}{95} = 0.0737$
Exposure odds in cases	${}_{exposure}O_{cases}$	${}_{exposure}O_{cases} = \frac{7}{44} = 0.1591$
Disease odds in the non-exposed	${}_{disease}O_{E-}$	${}_{disease}O_{E-} = \frac{7}{7} = 1.000$
Disease odds in the exposed	${}_{disease}O_{E+}$	${}_{disease}O_{E+} = \frac{44}{95} = 0.4632$

The intermediate measures (frequency measures) - necessary elements for calculation of exposure odds ratio as a measure of association in exposure comparisons calculated on the basis of data presented in Table 7 – are presented in Table 8. For demonstrating the equality of exposure odds ratio to disease odds ratio, both, the exposure odds in cases and controls, as well as disease odds in exposed and non-exposed are presented.

Calculation of measures of association in exposure comparisons

As just mentioned, in this group of measures we will present only one measure, namely the exposure odds ratio. The exposure odds ratio calculated according to Equation 10 based on data presented in Tables 7 and 8 is as follows (Equation 33):

$$EOR = \frac{\frac{7}{44}}{\frac{7}{95}} = \frac{0.1591}{0.0737} = \frac{7 \times 95}{44 \times 7} = \frac{95}{44} = 2.1591 = \quad \textbf{Equation 33.}$$

$$= \frac{\frac{7}{7}}{\frac{44}{95}} = \frac{1}{0.4632} = \frac{7 \times 95}{44 \times 7} = \frac{95}{44} = 2.1591 = DOR$$

The odds of being exposed to hyperirritable uterus is in group of mothers that experienced preterm delivery (cases) 2.16 times higher than in the group of mothers that did not experienced preterm delivery (controls), indicating that the association between preterm delivery and preterm delivery is strong.

Case study 2: Measures of potential impact

Measures assessing impact of an exposure on disease occurrence

An incidence study of impact of hyperirritable uterus on premature delivery

For demonstrating the calculation process in measures assessing the impact of hazardous factors we will use the same data set as presented in demonstration of calculation of measures of association in incidence comparisons – "An incidence study of impact of hyperirritable uterus on a premature delivery" (Table 1). We will use intermediate measures (frequency measures), presented in Table 2, as well as additional intermediate measures for assessment of population based measures, presented in Table 9.

Table 9. The intermediate measures (frequency measures) - necessary elements for calculation of population based measures for assessing the impact of exposure to a hazardous factor on disease occurrence

Frequency measure	Notation	Calculation and value
Risk in the population	R_{pop}	$R_{pop} = \frac{51}{800} = 0.0638$
Proportion of exposed in the population	P_{E+}	$P_{E+} = \frac{37}{800} = 0.0463$

Calculation of measures of assessment of impact of hazardous factors

The procedures of calculation of attributable risks in exposed and in population, and attributable fractions in the exposed and in the population are as follows:

1. Attributable risk in the exposed

   The attributable risk in the exposed calculated according to Equation 11 based on data presented in Tables 1, 2 and 9 is (Equation 34):

$$AR = 0.1892 - 0.0577 = 0.1315 \qquad \textbf{Equation 34.}$$

The risk for preterm delivery in mothers suffering from the hyperirritable uterus is 0.1892 (18.92%). In absolute terms, 0.1315 (13.15%) of this risk could be attributed to the hyperirritable uterus. This is the portion of the risk of a preterm delivery in the exposed individuals that could be eliminated if the exposure were eliminated.

2. Attributable risk in population

   The population attributable risk calculated according to Equation 12 based on data presented in Tables 1, 2 and 9 is as follows (Equation 35):

$$PAR = 0.0638 - 0.0577 = 0.0061 \qquad \textbf{Equation 35.}$$

In absolute terms, out of 0.0638 (6.36%) of risk of preterm delivery in a population, 0.0061 (0.6%) could be attributed to hyperirritable uterus. This is the risk of a preterm delivery in the population that could be eliminated if the exposure were eliminated.

3. Attributable fraction/percent in the exposed

   The attributable fraction in the exposed calculated according to Equations 13 and 14 based on data presented in Tables 1, 2 and 9 is as follows (Equations 36 and 37):

$$AF = \frac{0.1892 - 0.0577}{0.1892} = \frac{0.1315}{0.1892} = 0.6952 = 69.52\% \qquad \textbf{Equation 36.}$$

$$AF = \frac{3.2807 - 1}{3.2807} = 0.6952 = 69.52\% \qquad \textbf{Equation 37.}$$

The second procedure involves the use of relative risk (Equation 27). This result indicates that 69.52% of the risk of preterm delivery in the exposed group could be attributed to hyperirritable uterus. In other words, if those mothers suffering from hyperirritable uterus are spared from this unfavourable phenomenon (e.g. pharmacological treatment, rest, etc.), their risk of preterm delivery would decrease by 0.1315, what would represent 69.52% reduction of their preterm delivery incidence.

4. Attributable fraction/percent in the population

   The population attributable fraction calculated according to Equations 15 and 16 based on data presented in Tables 1, 2 and 9 is as follows (Equations 38 and 39):

$$PAF = \frac{0.0638 - 0.0577}{0.0638} = 0.0954 = 9.54\% \qquad \textbf{Equation 38.}$$

$$PAF = \frac{0.0463 \times (3.2807 - 1)}{1 + [0.0463 \times (3.2807 - 1)]} = 0.0954 = 9.54\% \qquad \textbf{Equation 39.}$$

The second procedure involves the use of relative risk (Equation 27). This result indicates that 9.54% of the total risk of preterm delivery in the population (in total, 6.38%) (exposed and non-exposed) could be attributed to hyperirritable uterus. In other words, 9.45% of preterm deliveries in the population could be prevented if all exposure to hyperirritable uterus were eliminated (e.g. pharmacological treatment, rest, etc.), or a reduction of 0.6 new cases of preterm delivery per 100 population (exposed and non-exposed) is expected if none of pregnant women is suffering from hyperirritable uterus. Such a reduction represents a 9.45% reduction of the incidence of preterm delivery in the population.

An incidence study of impact of medication with iron supplements during pregnancy on low birth weight of newborns

In an ambidirectional cohort study, based on Perinatal Informational System of Slovenia (PISS) (27), data of 800 mothers and their newborns were additionally analyzed. The observed outcome this time was low birth weight of a newborn (in this analysis 2500 g or less). Let us suppose that in Slovenia we have a preventive programme for reducing iron deficiency during pregnancy. On an individual level with iron deficiency treatment during pregnancy, a better oxygenation of a fetus could be attained and consecutively better growth ("exposure" to a protective factor), what could result in higher birth weight of a newborn. On a population level with iron deficiency reduction programme during pregnancy, a lower frequency of low birth weight at birth could be attained. Thus, we related the observed outcome to a medication with iron supplements during pregnancy.

The results of this analysis are presented in Table 10.

The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 10 – are presented in Table 11.

Table 10. Frequency of preterm delivery in two groups of mothers according to medication with iron supplements during pregnancy based on Slovene PISS data (27)

	Presence of hyperirritable uterus during pregnancy		Total
Preterm delivery	**No**	**Yes**	
No	401	348	749
Yes	39	12	51
Total	440	360	800

Table 11. The intermediate measures (frequency measures) - necessary elements for calculation of measures for assessing the impact of exposure to a protective factor on disease occurrence calculated on the basis of data from Table 10

Frequency measure	Notation	Calculation and value
Risk in the non-exposed	R_{E-}	$R_{E-} = \frac{40}{440} = 0.0886$
Risk in the exposed	R_{E+}	$R_{E+} = \frac{11}{360} = 0.0333$
Relative risk	RR	$RR = \frac{0.0333}{0.0886} = 0.3761$
Risk in population	R_{pop}	$R_{pop} = \frac{51}{800} = 0.0638$
Proportion of exposed in population	P_{E+}	$P_{E+} = \frac{360}{800} = 0.4500$

Calculation of measures of assessment of impact of protective factors

The procedures of calculation of prevented fractions in the exposed and in the population are as follows:

1. Prevented fraction/percent in the exposed

The prevented fraction in the exposed calculated according to Equations 17 and 18 based on data presented in Tables 10 and 11 is as follows (Equations 40 and 41):

$$PF = \frac{0.0886 - 0.0333}{0.0886} = \frac{0.0553}{0.0886} = 0.6239 = 62.39\% \qquad \textbf{Equation 40.}$$

$$PF = 1 - 0.3761 = 0.6239 = 62.39\% \qquad \textbf{Equation 41.}$$

The results indicate that medication with iron supplements during pregnancy has reduced the risk of low birth weight of a newborn by 62.39% among treated pregnant women. The iron supplements were 62.39% efficacious.

6. Prevented fraction/percent in the population

The population prevented fraction calculated according to Equations 19 and 20 based on data presented in Tables 10 and 11 is as follows (Equations 42 and 43):

$$PPF = \frac{0.0886 - 0.0638}{0.0886} = \frac{0.0249}{0.0886} = 0.2808 = 28.08\% \qquad \textbf{Equation 42.}$$

$$PPF = 0.4500 \times (1 - 0.3761) = 0.2808 = 28.08\% \qquad \textbf{Equation 43.}$$

The results indicate that the preventive programme for reducing iron deficiency has reduced the risk of low birth weight of a newborn by 28.08% in the population of pregnant women as a whole. We might argue that the iron deficiency reduction programme during pregnancy was 28.08% effective.

Measures assessing impact of an intervention on disease occurrence change

An incidence study of impact of medication with iron supplements during pregnancy on low birth weight of newborns

For demonstrating the calculation process in measures assessing the impact of an intervention on disease occurrence change we will use the same data set: "An incidence study of impact of medication with iron supplement during pregnancy on low birth weight of newborns" (Table 10). The only difference is that we suppose now that our data originate from a preventive trial in which one group of pregnant women were treated with iron supplements and the other was a control group. In calculations we will in fact use the same intermediate measures (frequency measures) as presented in Table 11, but since the notation is now slightly different, these intermediate measures are presented again in Table 12 according to this different notation.

Table 12. The intermediate measures (frequency measures) - necessary elements for calculation of measures for assessing the impact of an intervention on disease occurrence change calculated on the basis of data from Table 10

Frequency measure	Notation	Calculation and value
Risk in the control group (non-exposed)	R_C	$R_{E-} = \frac{40}{440} = 0.0886$
Risk in the treated (experimental) group (exposed)	R_T	$R_{E+} = \frac{11}{360} = 0.0333$
Relative risk	RR	$RR = \frac{0.0333}{0.0886} = 0.3761$

Calculation of measures of assessment of impact of an intervention on disease occurrence change

The procedures of calculation of absolute risk reduction, relative risk reduction and number needed to treat are as follows:

1. Absolute risk reduction

   The absolute risk reduction calculated according to Equation 21 based on data presented in Tables 10 and 12 is as follows (Equation 44):

$$ARR = 0.0886 - 0.0333 = 0.0553 \qquad \textbf{Equation 44.}$$

Absolute risk reduction is just the absolute difference in risks for observed outcome between the control and the treatment group. It is a less intuitive measure to interpret than relative risk reduction is, and its main role is to be used in calculation of number needed to treat. However, if we multiply the result in our case, and we get an absolute risk reduction of 5.53%, we can interpret it as follows: for every 100 pregnant women enrolled in the treatment group, about 5.5 "bad" outcomes (deliveries of a low birth weight newborn) would be prevented.

2. Relative risk reduction

   The relative risk reduction calculated according to Equations 22 and 23 based on data presented in Tables 10 and 12 is as follows (Equations 45 and 46):

$$RRR = \frac{0.0886 - 0.0333}{0.0886} = \frac{0.0553}{0.0886} = 0.6239 = 62.39\% \qquad \textbf{Equation 45.}$$

$$RRR = 1 - 0.3761 = 0.6239 = 62.39\% \qquad \textbf{Equation 46.}$$

Relative risk reduction measures how much of the risk is reduced in the experimental (treated) group compared to a control group. In our example, the result is 62.39%. This means that low birth weight of a newborn was reduced by 62.4% in the treatment group compared with the control group without treatment with iron supplements.

3. Number needed to treat

   The number needed to treat calculated according to Equation 24 based on data presented in Tables 10 and 12 is as follows (Equation 47):

$$NNT = \frac{1}{0.0553} = 18.08 \qquad \textbf{Equation 47.}$$

The result of calculating this measure indicates that for every 18 pregnant women treated with iron supplements, one newborn with low birth weight would be prevented.

Exercises

Task 1

In a maternity hospital data on successive 800 deliveries were collected in an ambidirectional cohort study. The observed outcome was low birth weight of a newborn, which is defined as birth weight 2500 g or less. The exposure under observation is smoking of mothers during pregnancy. The results of this study are presented in Table 13. Please:

- make small groups of students (maximum three students in a group);
- with other students in your group discuss what kind of comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13;
- calculate and make interpretation of all these measures;
- make a short presentation;
- present results to other groups of students;
- discuss your results vs. results of other students.

Table 13. Frequency of low birth weight in newborns in two groups according to smoking of mothers during pregnancy, based on PISS data (27)

Low birth weight of a newborn	Exposure to smoking of mother during pregnancy		Total
	No	Yes	
No	558	191	749
Yes	35	16	51
Total	593	207	800

Task 2

In the same study, the same observed outcome was related to a different hazardous factor, this time being elevated blood pressure of mothers during pregnancy. The results of this study are presented in Table 14. Please:

- make small groups of students (maximum three students in a group);
- with other students in your group discuss what kind of comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13;
- calculate and make interpretation of all these measures;
- make a short presentation;
- present results to other groups of students;
- discuss your results vs. results of other students.

Table 14. Frequency of low birth weight in newborns in two groups according to elevated blood pressure in mothers during pregnancy, based on PISS data (27)

Low birth weight of a newborn	Elevated blood pressure in mother during pregnancy		Total
	No	Yes	
No	707	42	749
Yes	46	5	51
Total	753	47	800

Task 3

In a cross-sectional study, based on CINDI Health Monitor survey 2001, which aimed at assessing the prevalence of health behaviours (28), data of 9,034 adults were analyzed. The observed outcome was frequent perception of stress without or with poor coping mechanisms. We related this outcome to gender as a risk factor for unfavourable health behaviour (i.e. in this context in a role of "exposure"). The results of this study are presented in Table 15. Please:

- make small groups of students (maximum three students in a group);
- with other students in your group discuss what kind of comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13;
- calculate and make interpretation of all these measures;
- make a short presentation;
- present results to other groups of students;
- discuss your results vs. results of other students.

Table 15. Prevalence of stress in both gender groups in the adult population of Slovenia in 2001, based on CINDI Health Monitor Survey, Slovenia 2001 (28)

	Gender		Total
Smoking	**Males**	**Females**	
No	3,235	3,570	6,805
Yes	861	1,321	2,182
Total	4,096	4,891	8,987

Task 4

In a case-control study, outcomes of successive deliveries were followed-up in a given period of time in a selected maternity. In this period, 36 deliveries of babies with low birth weight were registered (cases). For each case, approximately three controls out of other deliveries (i.e. with normal birth weights) were selected. The research question was whether elevated blood pressure of mother during pregnancy was associated with low birth weight of newborns. Data of 153 mothers and their respective newborns were analyzed. The results of this analysis are presented in Table 16. Please:

- make small groups of students (maximum three students in a group);
- with other students in your group discuss what kind of comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13;
- calculate and make interpretation of all these measures;
- make a short presentation;
- present results to other groups of students;
- discuss your results vs. results of other students.

Table 16. Frequency of preterm delivery in two groups of mothers according to absence (controls) or presence (cases) of hyperirritable uterus during pregnancy, based on Slovene PISS data (27)

	Presence of elevated blood pressure in mother during pregnancy		Total
Low birth weight	**No (Non-exposed)**	**Yes (Exposed)**	
No (Controls)	69	48	117
Yes (Cases)	30	6	36
Total	99	54	153

References

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 5, 2013.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999.
4. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
5. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
6. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization; 1998.
7. Centers for Disease Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics. Atlanta: Centers for Disease Control and Prevention. Available from URL: www.cdc.gov/osels/scientific_edu/ss1978/SS1978.pdf. Accessed: June 5, 2013.
8. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
9. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maidenhead: Open University Press, McGraw-Hill Education; 2005.
10. Babus V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
11. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
12. Pearce N. A Short Introduction to Epidemiology. Second Edition. Occasional Report Series No 2. Wellington, NZ: Massey University Wellington Campus, Centre for Public Health Research; 2003.

13. Walter SD. Choice of effect measure for epidemiological data. J Clin Epidemiol. 2000;53:931-9.
14. Symons MJ, Moore DT. Hazard rate ratio and prospective epidemiological studies. J Clin Epimiol. 2002;55:893-9.
15. Pearce N. Effect measures in prevalence studies. Environ Health Perspect. 2004;112:1047-50.
16. Barratt A, Wyer PC, Hatala R, McGinn T, Dans AL, Keitz S, Moyer V, For GG; Evidence-Based Medicine Teaching Tips Working Group. Tips for learners of evidence-based medicine: 1. Relative risk reduction, absolute risk reduction and number needed to treat. Can Med Assoc J 2004; 171:353–358.
17. Heller RF, Dobson AJ, Attia J, Page J. Impact numbers: measures of risk factor impact on the whole population from case-control and cohort studies. J Epidemiol Community Health 2002; 56:606–610.
18. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
19. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
20. Kleinbaum DG, Klein M. Survival analysis. A self-learning text. Second edition. New York: Springer; 2005.
21. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.
22. Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 1989.
23. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1993.
24. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.
25. Schunemann HJ, Guyatt GH. Clinical epidemiology. In: Ahrens W, Pigeot I (Eds.). Handbook of epidemiology. Berlin: Springer; 2005. p.1170-223.
26. Spasoff RA. Epidemiologic methods for health policy. New York: Oxford University Press, 1999.
27. Institute of Public Health of the Republic of Slovenia. Perinatal Informational system of the Republic of Slovenia. Available from URL: http://www.ivz.si/Mp.aspx?ni=46&pi=5&_5_id=56&_5_PageIndex=1&_5_groupId=185&_5_newsCategory=&_5_action=ShowNewsFull&pl=46-5.0. June 5, 2013.
28. Zaletel-Kragelj L, Fras Z, Maucec Zakotnik J (Eds.). Risky behaviours related to health and selected health conditions in adult population of Slovenia: results of Slovenia CINDI Health Monitor Survey 2001. II. Risky behaviours (in Slovene). Ljubljana: CINDI Slovenija, 2004. p.107-564.

Recommended readings

1. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
2. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999.
3. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
4. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
5. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
6. Schunemann HJ, Guyatt GH. Clinical epidemiology. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.1170-223.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Absolute risk assessment: From theory to practice**
Module: 2.9	ECTS (suggested): 0.25
Author(s), degrees, institution(s)	Mariana Dyakova, MD, PhD, Clinical Lecturer in public health Warwick Medical School, University of Warwick Emiliya Karaslavova, PhD, Assistant professor University of Plovdiv "Paisii Hilendarski", Plovdiv Elena Shipkovenska, MD, PhD, Professor Faculty of Public Health, Medical University of Sofia
Address for correspondence	Mariana Dyakova, MD, PhD, Clinical Lecturer in public health Warwick Medical School, University of Warwick Coventry, CV4 7AL E-mail: m.dyakova@warwick.ac.uk
Keywords	Absolute risk, cardiovascular diseases, chronic diseases, primary prevention, risk assessment, risk factor.
Learning objectives	After completing this module students should: • Know the characteristics and use of global (total/absolute) risk; • Be familiar with different methods of total risk assessment; • Understand the advantages and disadvantages of total risk assessment; • Be familiar with the specific implementation of global risk assessment for chronic non-communicable disease prevention and treatment; • Be aware of European populations' specificities and recommendations for total risk assessment of CVD; and • Be able to analyse the results from total risk assessment and consider it in health planning and management.
Abstract	Different risks to health interact and determine the multiple causality of disease. It is the combination and interaction of multiple causes that determines the absolute risk of any individual, i.e. the probability of event occurrence in a defined period of time. The concept of *"global/total risk" assessment* has been introduced, as a measure for individual absolute risk, as multiple risk factors confer greater risk than the sum of their components. The individual global risk assessment and the following risk stratification of a population appear to be relatively simple and cost-effective. On an individual level, it could be very effective for outcome prognosis and setting treatment objectives. On a population level, it offers opportunities to tailor prevention. Estimating the global/absolute risk for cardiovascular diseases (CVD) is the most widely used example of total risk assessment. The different methods (algorithms) for CVD total risk estimation are discussed and the specific example of the European SCORE implementation in the Bulgarian population is provided.
Teaching methods	An introductory lecture gives the students first insight in characteristics of global/absolute risk and its assessment. A number of case studies are presented, illustrating the practical use of different risk calculators. Students are asked to find other published studies, analyze and compare them through individual and group work.
Specific recommendations for teachers	• Work under teacher supervision/individual students' work proportion: 20%/80%; • Equipment: computers, LCD projector, internet, access to bibliographic data-bases; • Target audience: master degree students in public health/health care sciences; as CPD for primary care physicians.
Assessment of students	Design and presentation of a model for epidemiological study, involving total risk assessment.

ABSOLUTE RISK ASSESSMENT: FROM THEORY TO PRACTICE

Mariana Dyakova, Emiliya Karaslavova, Elena Shipkovenska

Theoretical background

Multifactorial causality and total risk assessment

Preventing death, disease, and injury requires systematic assessment and reduction of their determinants e.g. focusing on risks to health. This field has grown rapidly, focusing on the identification, quantification and characterization of threats to human health and the environment – a set of activities broadly called risk assessment (1). Governments need information from risk assessments that are comprehensive, reliable, relevant and timely to ensure population health and wellbeing. However, such information, which is crucial to prioritization and health policy, is typically limited. Many aspects should be considered in decision making to prevent/reduce risks to health (1): the extent of the threat posed by different risk factors, the availability of cost-effective interventions, and societal values and preferences are particularly important. Still most of the scientific efforts and healthcare resources are directed towards treating disease – the "rule of rescue" still dominates (2), while assessments of risk factors estimate the potential of prevention.

An effective risk assessment has a well-defined scope and considers a number of issues, such as: comprehensive risk factors identification; causal chain of events; various methodologies and health areas; type of determinant - outcome relationship; the life-course approach; timeliness and probability. Risk assessments typically use only attributable risk estimates, basically addressing the question: "What proportion of current burden is caused by the accumulated effects of all prior exposures?" Often a more policy-relevant question is: "What are the likely future effects of partial removal of current exposure?" Two key developments are therefore needed: an explicit focus on future effects and on less-than-complete risk factor changes (1).

Different health determinants interact to cause certain outcome (death, disease or injury) – *"a multiple causality of disease"* (4). The chain of events, leading to a specific outcome, includes proximal (direct), distal (indirect) factors and a number of intermediate ones. Separate estimation of the effects of individual risk factors does not typically take into account the effect of changes on the levels of other risk factors. Thus the *"causal web model of disease causation"* has been developed, reflecting the fact that risk factors often increase not only the risk of disease, but also levels of other risk factors (4). It is the combination and interaction of multiple causes that determines the absolute risk of an individual, i.e. the probability of event (disease) occurrence in a defined period of time. The absolute risk approach is counter-intuitive: e.g. benefit from lowering blood pressure does not depend (mainly) on level of blood pressure; it also differs from 'threshold' approaches of the past (5) and recommends that terms such as 'hypertension' and 'hypercholesterolaemia' should be abandoned. By a simple process of aggregation the same logic can be applied at a population level: in populations at high (average) absolute risk of a disease, all reversible risk factors should be lowered, and not because their risk factor levels are higher, but because the benefits of lowering them will be greater (6). Thus, the absolute risk approach is to be more extensively applied at both an individual (clinical) and population (public health) level.

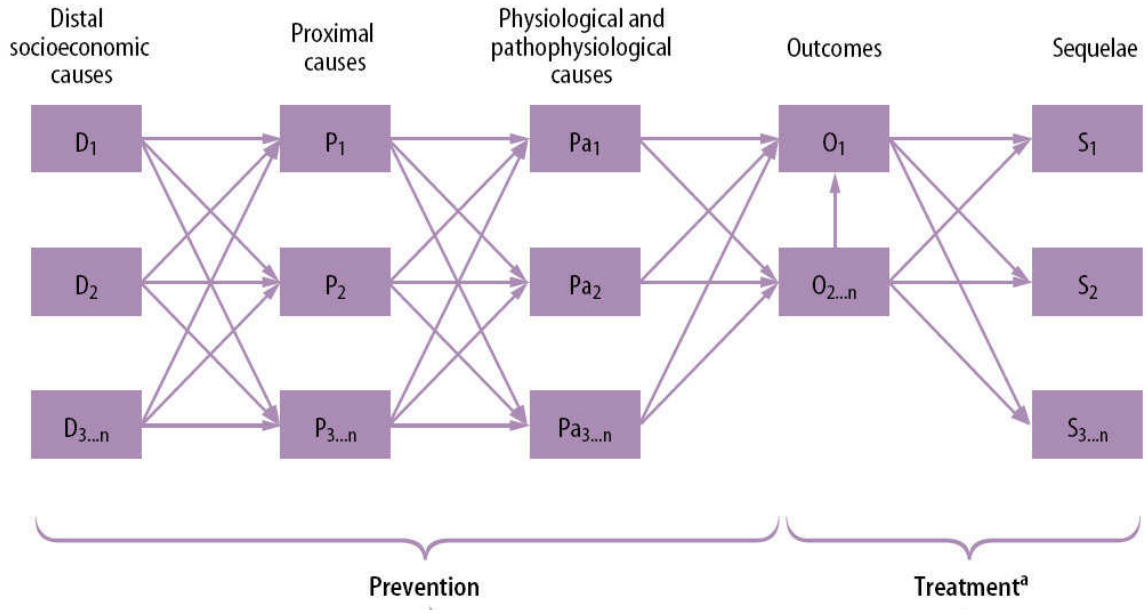
Multicausality and absolute risk assessment have important implications for prevention. The individual total risk assessment and the following risk stratification of a population (in low, moderate and high risk groups) appear to be relatively simple and cost-effective. On individual level it could be very effective for outcome prognosis and setting treatment objectives. On population level it offers opportunities to tailor prevention. The key message of multicausality is that different sets of interventions can produce the same goal, with the choice of intervention being determined by such considerations as cost, availability and preferences. Thus, prevention should not wait until further causes are elucidated; considerable gains can be achieved by reducing the risks to health that are already known (1).

Total risk assessment for chronic (cardiovascular) disease

A causal web model of a chronic non-infectious disease is given in figure 1. Distal socioeconomic causes include income, education, occupation, all of which affect levels of proximal factors such as inactivity, diet, tobacco use and alcohol intake; these interact with physiological and pathophysiological causes, such as blood pressure, cholesterol levels and glucose metabolism, to cause cardiovascular disease (CVD) such as coronary heart disease (CHD) or stroke. The sequelae include death or disability, such as angina or hemiplegia. Prevention of non-communicable diseases appears to be effective when it is timely; as much as possible risk factors are identified and modified (reduced); and risk factor reduction is as complete as possible and as long in life-course as possible. The proportion of a chronic disease risk that is attributable to a given risk factor is the proportion that would be avoided if lifetime exposure had been at a more favourable specified level (5,6). For binary harmful exposures, such as cigarette smoking, the attributable

disease burden in the population is estimated by comparing the actual burden with what would be expected in an 'unexposed' (non-smoking) population. However, for many important risk factors, zero exposure is not meaningful – e.g. blood pressure (BP), blood cholesterol (BChol), adiposity, etc. In these cases it is necessary to be explicit about the level (or distribution) of the risk factors with which the effects of the current levels (or distributions) are to be compared. That comparison level is referred to as the counterfactual level. Counterfactual exposure distributions can be more or less extreme in the contrast they provide with the actual distribution of exposure. The most extreme in the Burden of Disease approach is that distribution of a risk factor which confers theoretical minimum risk, i.e. "theoretical-minimum-risk exposure *distribution"* (7).

Figure 1. Causal chains of exposure leading to disease. Source: WHO World Health Report, 2002 (1)



Estimating the absolute risk for cardiovascular disease (CVD) is a typical and most widely discussed case of total risk assessment of non-communicable diseases. The CVD multicausal web has been extensively studied through the years and lots of the risk factors' interactions have been identified. Law and Wald have significant contribution to understanding how absolute risks of vascular disease should be managed in individuals (8,9):

1. Risk on a logarithmic (proportional) scale tends to be related linearly to the levels of continuously distributed risk factors (such as blood pressure and blood cholesterol), e.g. a 5 mmHg reduction in systolic BP results in constant percentage reduction in the risk of heart attack irrespective of the initial level of the BP.
2. Optimal values for vascular risk factors are well below current levels, so virtually everyone could benefit from lower values.
3. Some of the main determinants of CVD risk are not modifiable (age, sex and family history).

Thus, the intensity of efforts to reduce risk should depend on the absolute risk and not on the level of a particular risk factor; in individuals at high absolute risk all reversible risk factors should be lowered (9) – the absolute risk approach. This can be also extrapolated to population level: in populations at high (average) absolute risk of CVD, all reversible risk factors should be lowered. Patients are usually assessed and treated, according to the presence or absence of a particular risk factor (hypertension, obesity, diabetes, etc.). This approach can result in two unfavourable consequences: patients with low absolute cardiovascular risk are subjected to long-term (even life-long) treatment; patients with high absolute cardiovascular risk but without specific high level risk factor are left untreated. Another major disadvantage of the single-risk approach is that in most cases it is not cost-effective and thus it is inappropriate for low- and middle-resource settings. Current recommendations for the prevention of coronary heart disease (CHD) in clinical practice focus on total risk assessment.

Models of total risk assessment for CVD

The first coronary prediction algorithm published in 1991 is based on the famous cohort Framingham Heart Study (10,11), which followed 5209 adults in USA. In 1971 a second-generation group was enrolled - the Framingham Offspring Study and in 2005 the Third Generation Study was completed. Thus the absolute

risk formula has been periodically updated (12). The proposed risk function estimates the probability of developing CHD over the course of 10 years. A high risk is considered over 20% (10). European populations have also been affected by the CVD epidemic, mostly Scandinavian countries (13), as well as Central and Eastern European countries (14-16). Initially, Framingham data was used for several European risk prediction systems (17-19) and were incorporated into European guidelines, including the recommendations for CHD prevention in clinical practice by the European Societies on Coronary Prevention (20-22). However, a number of European studies revealed that risk calculators, based on the Framingham formula overestimate the total CVD risk in most European populations (23,24). Both countries with low CHD mortality and morbidity (e.g. Italy and France) and those with high CHD rates (e.g. Gemany and Norway) did not fit (25-27). The limitations of the Framingham risk calculator and its disadvantages for implementation in European populations are summarized below (28-31):

1. Framingham score sheets are only for people without known CHD.
2. The algorithm is only for coronary heart disease, not for other heart and vascular diseases.
3. It is based on a relatively small cohort, predominantly Caucasian.
4. Some of the sex-age groups are quite small, thus estimation may lack precision.
5. Framingham risk score is for a CVD risk in a 10-year period and may not adequately reflect the long-term/lifetime CHD risk of young adults.
6. The presence of any elevated CHD risk factor requires attention because it may confer a high risk in the long run, even if the 10-year risk does not appear to be high.
7. Framingham algorithm does not include important risk factors as family history / genetic predisposition, body mass index (BMI) and social status.
8. Framingham score may over-identify older people for aggressive interventions as age is increasing risk substantially, so relative risk may be better estimate for elderly.
9. Framingham Study (10) includes all non-fatal CHD end-points (e.g. unstable angina) in the formula, making it difficult to validate the function for other studies.
10. It is difficult to adjust the model with local data for use in other countries.
11. The Framingham score **should not be used** instead of a medical examination.
12. The Framingham cohort is representative of the high coronary risk American population from the 70^{es}/ 80^{es} of the 20^{th} Century. At present, CVD is declining in most developed economies, so this calculator tends to overestimate the up-to-date absolute risk.

One of the first and most widely used European based total risk prediction algorithms was PROCAM, named after the extensive German "PROspective CArdiovascular Münster (PROCAM) Study", started in 1978 (32,33). The results from the following PRIME Study (Prospective Epidemiological Study of Myocardial Infarction) showed that both Framingham and PROCSM risk calculators overestimate the absolute cardiovascular risk in some European populations (e.g. in UK and France) and recommended the elaboration of a separate national risk prediction scores (34). Such algorithms were made for the UK: separately for England - QRISK (Cardiovascular disease risk score for the United Kingdom: prospective open cohort study, 35) and Scotland - ASSIGN, based on the Scottish Heart Health Extended Cohort (36) and West of Scotland Coronary Prevention Study (37). The last one included also the impact of socio-economic status in the total risk prediction equation. The European Society of Cardiology and the Second Joint Task Force instigated the development of a risk estimation system based on a large pool of representative European data sets that would capture the regional variation in risk. This led to the establishment of the SCORE (Systematic COronary Risk Evaluation) project to develop a risk scoring system for use in European clinical practice in liaison with the Third Joint Task Force (31,38).

Characteristics, new aspects and advantages of the SCORE calculator (31, 39)

The SCORE project pooled data from **12 European countries cohort studies**. Most were population-based, though some occupational cohorts were included to increase representation of regions of lower risk. Subjects were excluded from the development of the risk chart if they had a previous history of heart attack.

The most important feature of the SCORE method is that it is estimating total cardiovascular risk rather than risk of CHD. This is a shift from the traditional epidemiological approach of aetiology to a public health perspective focusing on the consequences of risk factors. Non-coronary cardiovascular disease is important because it represents a greater proportion of all CVD risk in European regions with low rates of CHD.

The SCORE project shift the emphasis in risk estimation to fatal cardiovascular disease events only. There is no doubt that both patients and physicians are interested in non-fatal CVD events as well as they are a major economic burden for the health care system and the society. However, non-fatal CVD events are critically dependent on definitions and diagnostic methods. The SCORE project used 'hard' CVD end-points as its aim was to develop cardiovascular risk estimation systems applicable on national level. Many European countries do not have cohort studies on cardiovascular disease, but all countries have national cause specific mortality data, which can be used to estimate the baseline risk of the population.

Changing thresholds for high risk to 5% - shift in the risk estimation from any CHD event to the risk of fatal CVD end-point also means redefinition of the threshold for intensified risk modification efforts. To emphasise that there is no single level of absolute risk that defines an optimal threshold for risk factor intervention, regardless of individual age, sex or nationality, the SCORE risk charts display the 10-year risk of cardiovascular death both as figures as well as categories. Health economic research has suggested that risk threshold for cost effectiveness of risk factor interventions, such as cholesterol lowering drug therapy, is not a simple function of absolute risk but also varies with age and sex.

Versions for total cholesterol and cholesterol/HDL ratio. Persons with multiple risk factors tend to have lower HDL cholesterol levels and there is therefore a concern that failing to take HDL cholesterol into account will underestimate risk in those most at risk.

Versions for low and high cardiovascular risk European populations - two SCORE charts were developed – one for low CVD risk, including France, Belgium, Italy etc and one for high CVD risk, including Scandinavian countries, Germany etc.

Change in the ages for which the risk is displayed. SCORE risk charts are providing more detail in the age group 50–65 years where risk changes most rapidly. Risk for age 30 has been suppressed as in many SCORE datasets there were no events in this age group.

SCORE risk charts are intended for risk stratification in the primary prevention of CVD. There are no risk estimates for CHD patients, as there is now a consensus that these people should be treated as high risk cases (38).

Limitations of the SCORE study and risk calculator

The underlying risk functions are based on single risk factor measurements, not on the persons "usual" levels. The charts also consider only the principal risk factors. Data for the SCORE project was gathered during the 70^{ies} and 80^{ies} of the 20th Century. CVD death rates are decreasing with 30% to 50% on average in Western Europe (40). It is highly probable that SCORE calculator predicts higher absolute risk, e.g. a total CVD risk of 5% in 1985 could be equal to 2,5% in 2003 (41). Uncritical application of any risk score may mislead patients and health professionals and ongoing studies are needed to ensure CVD risk assessment is accurate and up-to-date (42).

CVD risk assessment to low-resource settings (43)

The most common application of risk stratification is as a tool for cost-effective health policy decisions. Given a level of health expenditure, risk stratification can help identify the subset of patients most in need of treatment. It is important in this context, however, not to confuse this triage application with a cost-effectiveness analysis. As a result of a formal cost-effectiveness study, the group of patients that would benefit from treatment might be much larger than those whom current budgets can accommodate. It is imperative therefore, that the necessary epidemiological data be obtained so that future decisions can be based on evidence. Feasible risk-assessment methods using simple clinical indicators that are measurable in less well-resourced settings can be used to develop a pragmatic risk-stratification system. Such systems, although less accurate, are likely to be the only feasible option in such contexts (43).

CVD risk scoring validation, application and limitations

Much research has been undertaken to validate different CVD risk scoring methods, so that individual CVD risk is correctly identified (44,45). Commonly used risk functions have been seen to estimate absolute risk poorly in some ethnic groups, low-risk groups, and high-risk groups including people with diabetes (46). Modifications to existing scores and new methods developed in appropriate populations have aimed to improve the accuracy of cardiovascular risk estimation. As with any new technology, risk functions and scores must be shown to have a favourable influence on cardiovascular outcomes, patient's risk factors, risk behaviours or treatment if they are to be used in a primary preventive strategy. Previous evidence from diverse areas of healthcare suggest that structured clinical decision aids have only limited value, with most encouraging benefits in relation to attitudes and knowledge, but with limited effects on health (47). Whatever scoring mechanism is used, assessing someone's level of CVD risk does not actually change it. Short emphasizes that there is no advantage in assessment, without the ability to intervene and to make changes to lower that risk (48). A significant proportion of CVD morbidity and mortality can be prevented through population strategies for primary prevention. Efficient and effective means of identifying high-risk individuals and then providing the support to enable them to modify their lifestyles requires a delivery system which gives priority to preventive services rather than focusing on treatment (49). Despite various public health and clinical efforts for primary prevention of CVD, a large number of the population, considered at increased risk of vascular disease, remains unidentified, untreated and not reached by lifestyle advice or intervention (50).

Case study

Absolute cardiovascular risk assessment in the Bulgarian urban population

Rationale

Cardiovascular diseases are a major problem for the Bulgarian population, accounting for more than 60% of the all-cause mortality and a considerable percent of the disability in the last two decades (51). Ineffective primary prevention of CVD and delayed emergency care are considered the main factors for the high cardiovascular morbidity and mortality.

Absolute risk assessment for CVD and its use in everyday clinical practice has been widely discussed among Bulgarian epidemiologists and recommended by cardiologists, in response to the European guidelines. Some research studies attempted to adapt the SCORE calculation charts to the Bulgarian demographics, but the results were not applied into clinical practice, neither were used for health policy recommendations. An extensive, up-to-date population-wide research, aimed at the absolute CVD risk assessment and development of nationally relevant cardiovascular prevention guidelines for clinical practice was necessary.

The case study presented here is a brief summary of the methods and results from the largest most recent cross-sectional epidemiological research, exploring the SCORE implementation for absolute risk assessment in the Bulgarian urban population.

Study aim and methodology

A cross-sectional observation study was conducted during the period 2005-2007.

The study objectives were:

1. Overview of the main risk factor burden for CVD in Bulgaria (healthy population);
2. Evaluation of the absolute (total) cardiovascular risk, using the European SCORE calculator and population stratification for clinical and health policy recommendations.

Study population: a representative sample for the Bulgarian urban population, 3810 subjects in total, age 25-74, men and women, was investigated. The study was conducted in the five biggest Bulgarian cities with the help of 106 general practitioners (GPs). The sample was randomly selected, with no history of fatal cardiovascular event or serious CVD.

Risk factors studied: age, sex, body mass index (BMI), blood pressure, waist and hip circumference, serum lipids, glucose, profession, education, diet, physical activity, behavioural characteristics etc.

The absolute 10-year risk of a fatal CVD event was estimated using the SCORE (HeartScore®) high risk chart (31) and the European guidelines, 2003 (38). A risk level categorization of the population sample was done, defining groups with low, intermediate and high risk. The data was compared to other similar European studies. A relevant threshold for a high-risk prevention strategy was proposed. The total risk assessment was performed through a specifically developed software application, which consisted of a CVD Programme for global risk estimation ("CardioDB") and "Scorecard Ver 4.0.0.15" (electronic version of SCORE charts). Each general practitioner received this application, where he/she filled in the required data as well as calculated the actual absolute risk of the individual; this also created a database for future medical check-ups. Population stratification was done, according to the European guidelines, 2003 (38):

Low risk group - SCORE ≤ 1%

Intermediate risk group - SCORE between 2% and 4%

High risk group - SCORE ≥ 5%

Results[8]

CVD risk factor and SCORE distribution in the studied population is presented in tables 1 and 2. The data reveal that the mean and median values of most of the studies CVD risk factors are slightly increased or borderline, compared to clinical thresholds. This is in contrast with the considerably high total risk score, which is over 5% for the total sample, studied and have very high mean value in men – 9.12%. However the results for SCORE correspond to the high CVD death rate in the Bulgarian population, which once again is conclusive that not individual risk factor levels are to be examined and treated, but the total cardiovascular risk. Another interesting result is the considerably higher average risk score for men then fro women, which also cannot be explained by the individual risk factor levels.

[8] Here, only limited data is presented for illustration of the theoretical background.

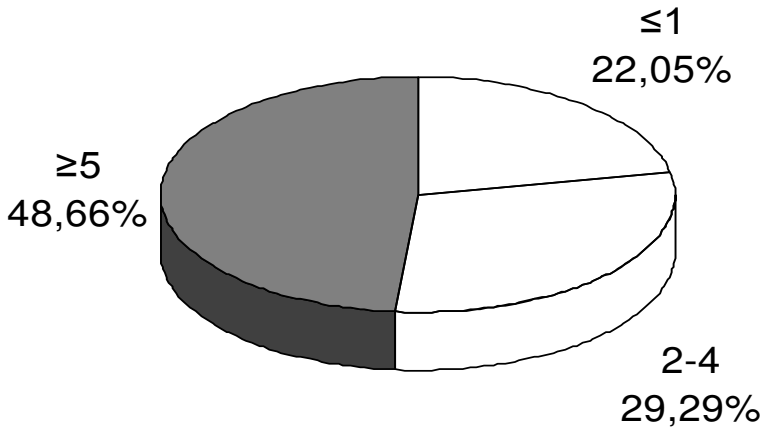
Table 1. CVD risk factors' and SCORE median, mean and standard deviation in the sample

CVD risk factors	Median	X (Mean)	σ (SD)
Age	59	58,31	9,25
Body mass index – BMI	27,02	27,67	4,49
Systolic blood pressure – SBP	140	139,23	20,03
Diastolic blood pressure – DBP	82	86,36	11,97
Total Cholesterol	5,10	5,67	1,47
HDL cholesterol	1,10	1,12	0,36
LDL cholesterol	3,40	3,46	1,06
SCORE	4,00	6,43	6,62

Table 2. CVD risk factors' and SCORE mean and standard deviation, according to sex

Risk factors	MEN		WOMEN	
	Mean	SD	Mean	SD
Age	56,90	9,81	59,47	8,59
BMI	27,99	3,87	27,41	4,93
SBP	152,41	21,36	147,72	20,51
DBP	92,16	12,11	88,87	11,65
Total Cholesterol	5,63	1,38	5,70	1,54
LDL	3,43	1,07	3,50	1,06
HDL	1,13	0,38	1,11	0,96
SCORE	**9,12**	**8,08**	**4,21**	**3,89**

Figure 2. SCORE stratification of the studied population, according to ESC, 2003



According to figure 2, almost 50% of the studied people had SCORE>5%, i.e. nearly half of the sample belongs to the "high risk group" for fatal cardiovascular event in the next 10 years.

In relation to the above results and considering the lack of a national-specific absolute cardiovascular risk calculator, it was of interest to make a further categorization of the "high risk group" into "relatively high", "very high" and "excessively high risk":

Relatively high risk for fatal CVD - SCORE ≥ 5%

Very high risk for fatal CVD - SCORE ≥ 10%

Excessive risk for fatal CVD - SCORE ≥ 15%

Table 3. "High risk" distribution of the sample population

SCORE	Number	%
5-9	927	24,31
10-14	512	13,46
≥15	415	10,89
Total	3810	100,0

Table 3 shows that nearly one quarter of the studied sample has an absolute CVD risk, according to SCORE of over 10% (13,46 + 10,89 = 24,35%). Even this is an economically not cost-effective and not applicable to Bulgarian healthcare system and resources threshold for high-risk prevention strategy group. The people with SCORE>15% are nearly 10% of the studied population, which appears to be already a reasonable group for high-risk strategy for CVD primary prevention.

The age and sex distribution of the sample population with SCORE over 5% is presented in figures 3 and 4. Most of the high risk subjects are found in the age groups over 45 years and especially over 55. The percent of men with SCORE>5% is significantly higher than that of women ($p<0.05$).

Figure 3. Age distribution of "high risk group" (SCORE>5%)

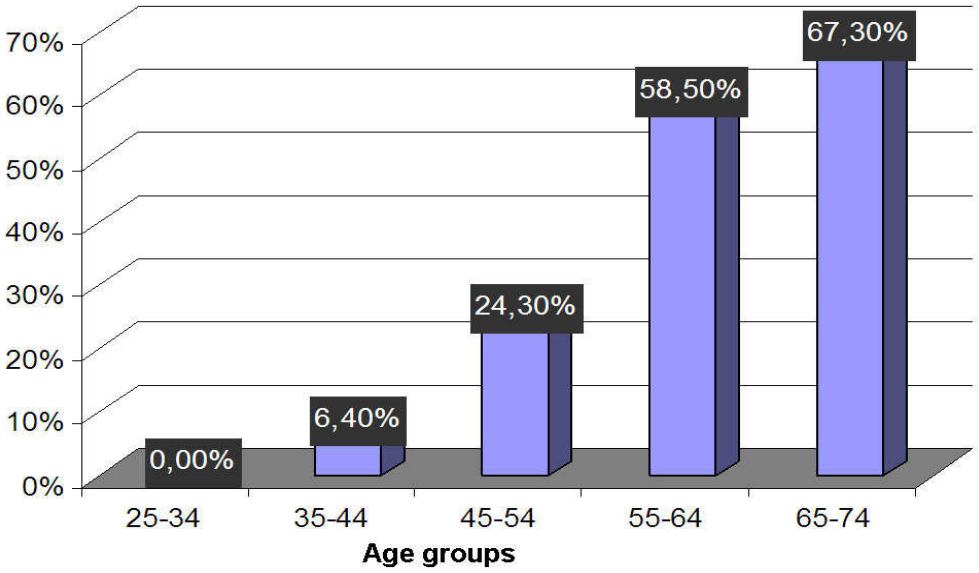
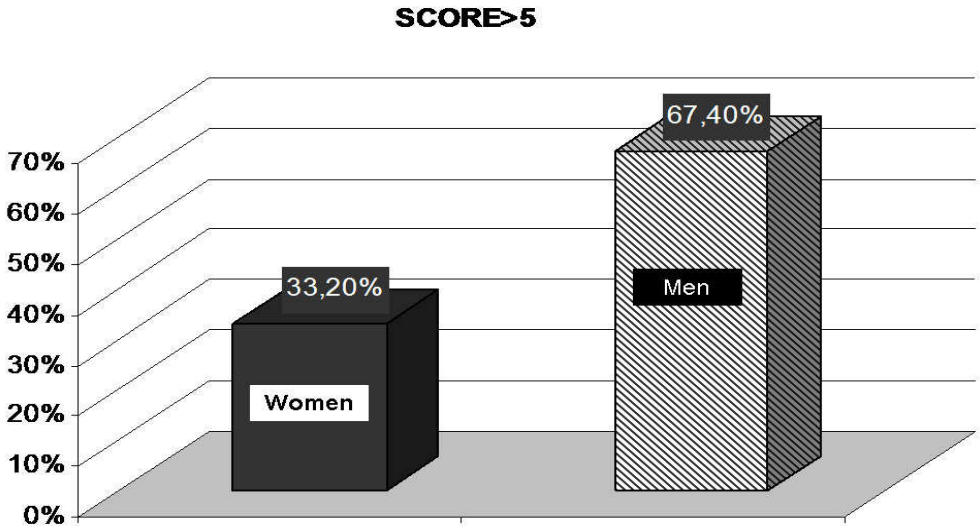


Figure 4. Sex distribution of "high risk group" (SCORE>5%)



The SCORE stratification on figures 5 and 6 is characterized by:

1. Low absolute cardiovascular risk in the age group 25-34 in both men and women - 100% of the sample has SCORE<1%. However this could be explained by the small number of examined subjects in this group.
2. The SCORE (total risk) difference between men and women in the age groups 35-44 and 45-54 is significant ($p<0.05$). Women maintain higher percent with SCORE<1% in both age groups, while the percent of men with SCORE over 2 is rising. Nearly half of the men aged 45-54 years (42,4%) have SCORE>5%.
3. The gender differences in the age groups 55-64 and 65-74 are increasing.
4. The global CVD risk is increasing steeply in the age groups over 45. Only 6,6% of the men, aged 65-74 have low and average absolute risk (SCORE<5%), while the percentage of excessively high risk (SCORE>15%) reaches 46,6%, i.e. nearly half of the men over 65 years of age have extremely high absolute risk of developing a fatal cardiovascular event in the next 10 years.

Figure 5. SCORE groups distribution according to age – women

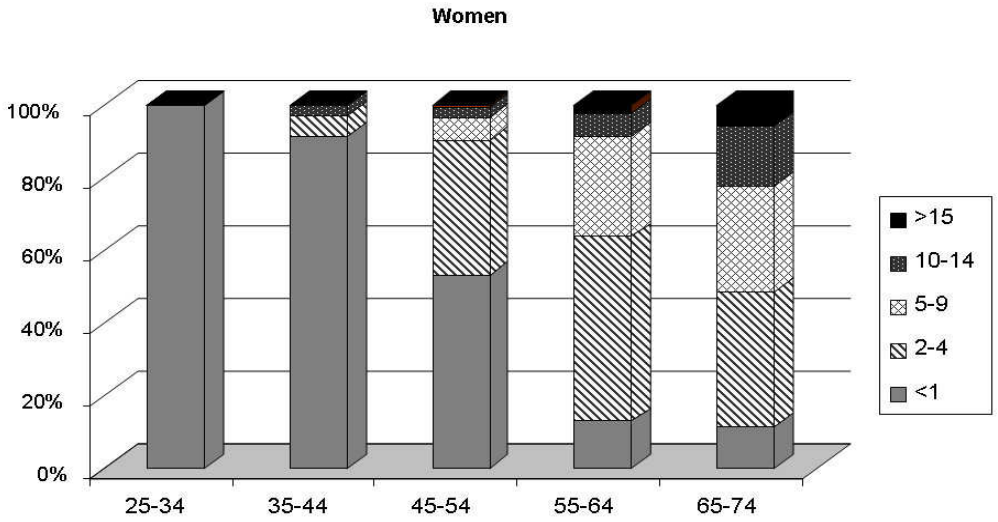
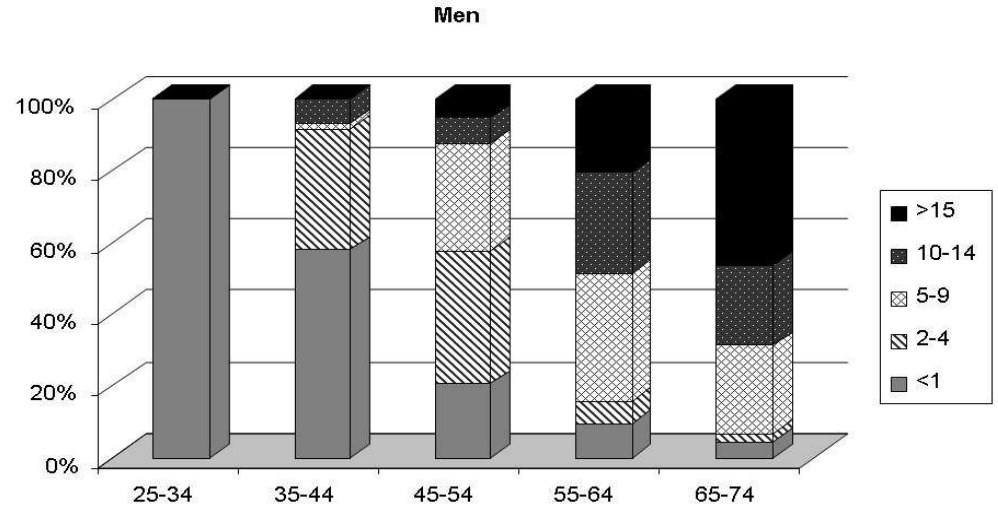


Figure 6. SCORE groups distribution according to age – men



Discussion

The high percentage of the studied population with SCORE>5% is consistent with the high CVD death rate in the Bulgarian population. It can be also considered as an indicator for a present "hidden morbidity", the so-called "morbidity iceberg", as most of the examined subjects are free of serious CVD, have relatively low levels of the main studied cardiovascular risk factors and most of them do not receive any medication treatment.

One of the main advantages of the SCORE algorithm and the ESC guidelines, 2003 is the possibility for national-specific adaptation of the risk calculator and fixing a national relevant threshold for CVD high risk primary prevention strategy. However, the SCORE project considers an absolute risk of over 5% as high, but that is not the only element, on which health policy decision for CVD prevention targets should be taken. It has to take into account the health care resources available in the system as well as the general socio-economic, environmental and cultural background. Thus, according to our data, the most possibly cost-effective threshold for intensive risk factor modification strategy could be 15%.

According to the 2003 European guidelines on cardiovascular disease prevention in clinical practice (38), the SCORE method is leading in assessing global (absolute) cardiovascular risk. They set three specific objectives, concerning CVD prevention:

1. Adaptation of the proposed guidelines and recommendation at national level;
2. A tool for prioritizing patients (i.e. population stratification); and
3. A tool for counselling in clinical practice.

Our findings can be discussed with regard to these three explicit objectives of the guidelines:

Adaptation to national specificity. The regional and local differences in morbidity and mortality as well as different population risk profiles require risk evaluation and scoring systems development against epidemiological data from the target population to be screened before implementation in clinical practice. Whether the high risk chart applies to Bulgaria still needs more investigation and comprehensive population-representative studies.

Tool for prioritising patients. Our study found that the guidelines are unlikely to serve as an effective tool for prioritising Bulgarian population, as they classify an unreasonable number of people as at high risk (over 5%).

Tool for counselling in clinical practice. The European guidelines for CVD prevention are clearly recommended for direct use in counselling in clinical practice (38). Several ethical dilemmas arise from the likelihood of overestimating someone's true risk for cardiovascular disease (52). The systematic coronary risk evaluation project does not discuss the problem of retrospective risk bias. We question whether it was scientifically justifiable to include the risk charts of the systematic coronary risk evaluation project in guidelines intended for implementation in a clinical setting before validation in a contemporary context. Any overestimation of a person's risk for cardiovascular disease can have important implications. Apart from causing unnecessary concern, it undermines the patient's informed choice for intervention. It is also likely to increase prescribing costs and affect life insurance premiums. As yet little scientific knowledge is available on how the communication of this kind of risk affects people's understanding of themselves, their bodies, and their lives.

Conclusion

Despite the contribution of numerous experts and professional societies, it seems that authoritative clinical guidelines on the basis of the systematic coronary risk evaluation project may be an example of premature application of medical technology in routine clinical practice. On the other hand, the CVD epidemic in developing countries requires an immediate action and effective solution. The insufficiency of enough local/national data for development of a specific population targeted risk assessment tool should not be an excuse for the lack of public health inertia and health policy inactivity, addressing the problem.

Exercises

Task 1: Search for risk factors for a socially-important chronic/non-communicable condition (e.g. diabetes, hypertension, osteoporosis etc) and build a multicausal web for it.

Task 2: Find, choose and learn about two total risk scores/calculators for estimating cardiovascular risk; design a hypothetical epidemiological study to compare their applicability and validity for a country of your choice. Use the other methodological modules to help you.

References

1. World Health Organization. The World Health Report 2002: Reducing risk, promoting healthy life. Geneva: WHO, 2002:9–26. Available at: http:// www.who.int/ whr/2002/en/ index.html. (Accessed 03 June, 2008).
2. Hope T. Rationing and life-saving treatments: should identifiable patients have higher priority? Journal of Medical Ethics 2001;3:179-85.
3. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. International Journal of Epidemiology 2002; 31:285-93.

4. Rothman KJ, Greenland S. Modern epidemiology. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1998.
5. Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. Am J Epidemiol. 1974;99:325–32.
6. Walter SD. Prevention for multifactorial diseases. Am J Epidemiol. 1980;112:409–16.
7. Murray CJ, Lopez AD. On the comparable quantification of health risks: lessons from the Global Burden of Disease Study. Epidemiology. 1999;10:594–605.
8. Wald NJ, Law MR. A strategy to reduce cardiovascular disease by more than 80%. BMJ 2003;326:1419-23.
9. Law MR, Wald NJ. Risk factor thresholds: their existence under scrutiny. BMJ 2002;324(7353):1570-6.
10. Wang TJ et al. A Risk Score for Predicting Stroke or Death for Individuals in the Community: The Framingham Heart Study. JAMA 2003;290:1049-56.
11. Kannel W. The Framingham Experience. Am. J. Car. 1976;77:6-10.
12. Kannel WB, Feinleib M, McNamara PM et al. An investigation of coronary heart disease in families. The Framingham offspring study. Am J Epidemiol 1979;110:281–90.
13. Kastarinen MJ, Saloma VV, Vartiainen EA, et al. Trend in blood pressure levels and control of hypertension in Finland from 1982–1997. J Hypertens 1998;16:1379–87.
14. Shapo L, Pomerleau J, McKee M. Epidemiology of hypertension and associated cardiovascular risk factors in a country in transition: a population based survey in Tirana City, Albania. J Epidemiol Community Health 2003;57:734–9.
15. Kuulasmaa K, Tunstall-Pedoe H, Dobson A, et al. Estimation of contribution of changes in classic risk factors to trends in coronary-event rate across the WHO MONICA Project population. Lancet 2000;3555:675–87.
16. Powles JW, Zatonski W, Vander Hoom S, Ezzati M. The contribution of leading diseases and risk factors to excess losses of healthy life in Eastern Europe: burden of disease study. BMC Public Health 2005;5:116.
17. Schulte H, Assmann G. CHD risk equations obtained from the Framingham Heart Study and applied to PROCAM study. Cardiovasc Risk Factors 1991;1:126–33.
18. Haq IU, Ramsay LE, Yeo WW et al. Is the Framingham risk function valid for northern European populations? A comparison of methods for estimating absolute coronary risk in high risk men. Heart 1999;81:40–6.
19. Marrugat J, D'Agostino R, Sullivan L, et al. An adaptation of the Framingham coronary heart disease risk function to European Mediterranean areas. J Epidemiol Community Health 2003;57:634–8.
20. Pyorala K, De Baker G, Graham I, Poole-Wilson and Wood D. Prevention of coronary heart disease in clinical practice. Recommendations of the Task Force of the European Society of Cardiology, European Atherosclerosis Society and European Society of Hypertension. Eur Heart J 1994;15:1300–31.
21. Wood DA, De Backer G, Faergeman O et al. Prevention of coronary heart disease in clinical practice. Recommendations of the Second Joint Task Force of European and other Societies on coronary prevention. Eur Heart J 1998;19:1434–503.
22. Wood D, De Backer G, Faergeman O et al. Prevention of coronary heart disease in clinical practice. Summary of recommendations. Coronary Prevention. Blood Press 1998;7:262–9.
23. Ducimetière P, Richard JL, Cambien F et al. Coronary heart disease in middle-aged Frenchmen. Comparisons between Paris Prospective Study, Seven Countries Study, and Pooling Project. Lancet 1980;1:1346–50.
24. Menotti A, Puddu PE, Lanti M. Comparison of the Framingham risk function-based coronary chart with risk function from an Italian population study. Eur Heart J 2000;21:365–70.
25. Hense HW, Schulte H, Lowel H et al. Framingham risk function overestimates risk of coronary heart disease in men and women from Germany-results from the MONICA Augsburg and the PROCAM cohorts. Eur Heart J 2003;24:937–45.
26. Brindle P, Emberson J, Lampe F, Walker M, Whincup P, Fahey T, et al. Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. BMJ 2003;327:1267-72.

27. Van Steenkiste B, van der Weijden T, Stoffers HEJH, Grol R. Barriers to implementing cardiovascular risk tables in routine general practice. Scand J Prim Health Care 2004;22:32-7.
28. Brindle P, Beswick A, Fahey T, Ebrahim S. Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. Heart 2006;92:1752-9.
29. Jones AF, Walker J, Jewkes C et al. Comparative accuracy of cardiovascular risk prediction methods in primary care patients. Heart 2001;85:37–43.
30. Wilson PWF, D'Agostino R, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837-47.
31. Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J 2003;24:987-1003.
32. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study. Circulation 2002;105:310–5.
33. Assmann G, Cullen P, Schulte H. The Münster Heart Study (PROCAM). Results of follow-up at 8 years. Eur Heart J 1998; 19 (Suppl A):A2-A11.
34. Empanaa JP et al. (on behalf of the PRIME Study Group) Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME Study. European Heart Journal 2003;24:1903–11.
35. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK. A new cardiovascular disease risk score for the UK. BMJ 2007;335:136.
36. Haq IU, Jackson PR, Yeo WW, Ramsay LE. Sheffield risk and treatment table for cholesterol lowering for primary prevention of coronary heart disease. Lancet 1995;346:1467-71.
37. The West of Scotland Coronary Prevention Study Group. West of Scotland Coronary Prevention Study. Baseline risk factors and their association with outcome in the West of Scotland Coronary Prevention Study. Am J Cardiol 1997;79:759–62.
38. De Backer G, Ambrosioni E, Borch-Johnsen K, et al. European guidelines on cardiovascular disease prevention in clinical practice. Third Joint Task Force of European and other Societies on Cardiovascular Disease Prevention in Clinical Practice. Eur J Cardiovasc Prev Rehabil 2003;10(Suppl 1):S1–78.
39. Cowie M. R. Coronary risk - time for a more sophisticated approach? European Heart Journal 2002;23:589–591.
40. Kuulasmaa K, Tunstall-Pedoe H, Dobson A, et al. Estimation of contribution of changes in classic risk factors to trends in coronary-event rate across the WHO MONICA Project population. Lancet 2000;3555:675–87.
41. Holmen J, Midthjell K, Krьger Ц, Langhammer A, Holmen TL, Bratberg GH, et al. The Nord-Trцndelag health study 1995-7 (HUNT-2): objectives, contents, methods and participation. Norsk Epidemiologi 2003;13:19-32.
42. Vishnu Madhok and Tom Fahey. Cardiovascular risk estimation: important but may be inaccurate. BMJ 2006;332;1422. Available at: http://bmj.com/cgi/content/full/332/7555/1422 (Accessed August 20, 2013).
43. World Health Organization. Integrated management of cardiovascular risk: Report of a WHO meeting; 9-12 July 2002, Geneva: WHO, 2002.
44. Brindle P, Beswick A, Fahey T, Ebrahim S. Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. Heart 2006;92(12):1752–9.
45. Beswick AD, Brindle P, Fahey T, Ebrahim S. A Systematic Review of Risk Scoring Methods and Clinical Decision Aids Used in the Primary Prevention of Coronary Heart Disease (Suppl). NICE Clinical Guidelines No. 67S. Available from:

    http://www.ncbi.nlm.nih.gov/books/NBK55818/. London: Royal College of General Practitioners, 2008.
46. Beswick A, Brindle P. Risk scoring in the assessment of cardiovascular risk. Current Opinion in Lipidology 2006;17(4):375–86.
47. Estabrooks C, Goel V, Thiel E, Pinfold P, Sawka C, Williams I. Decision aids: are they worth it? A systematic review. Journal of Health Services & Research Policy 2001;6(3):170–82.
48. Short R. Putting Vascular Disease Management into Practice. Bath: Medical Management Services, 2009.

49. Bernard SL, Lux L, Lohr KN (RTI International). QQUIPP: Healthcare Delivery Models for Prevention of Cardiovascular Disease (CVD). London: The Health Foundation, 2009.
50. Dyakova M, Drew C, Wright N, Clarke A, Rees K. Systematic versus opportunistic risk assessment for the primary prevention of cardiovascular disease. Cochrane Database of Systematic Reviews 2013, Issue 2. Art. No.: CD010411.
51. National Statistical Institute, 2008. Available at:
http://www.nsi.bg/SocialActivities/Health.htm (Accessed May 28, 2013).
52. Getz L, Kirkengen AL, Hetlevik I, Romundstad S, Sigurdsson JA. Ethical dilemmas arising from implementation of the European guidelines on cardiovascular disease prevention in clinical practice: descriptive epidemiological study. Scand J Prim Health Care 2004;22:202-8.

Recommended readings

1. Lopez AD, Mathers CD, Ezzati M et al. Measuring the Global Burden of Disease and Risk Factors, 1990—2001. Global Burden of Disease and Risk Factors,ed. 1-13. New York: Oxford University Press, 2006.
2. Reynolds TM, Twomey PJ, Wierzbicki AS. Concordance evaluation of coronary risk scores: implications for cardiovascular risk screening. Curr Med Res Opin 2004;20:811-8.
3. Marteau TM, Kinmonth AL. Screening for cardiovascular risk: public health imperative or matter for individual informed choice? BMJ 2002;325:78-80.
4. Saltman RB, Figueras J. European Health Care Reform: Analysis of Current Strategies. Copenhagen: WHO Regional Office for Europe, 1997.
5. World Health Organization. Integrated management of cardiovascular risk: Report of a WHO meeting; 9-12 July 2002, Geneva: WHO, 2002.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Introduction to epidemiological studies**
Module: 2.10	ECTS (suggested): 0.25
Author(s), degrees, institution(s)	Enver Roshi, MD, PhD, Professor – Faculty of Public Health, University of Medicine, Tirana, Albania; Genc Burazeri, MD, PhD, Professor – Faculty of Public Health, University of Medicine, Tirana, Albania; Ervin Toci, MD – Faculty of Public Health, University of Medicine, Tirana, Albania.
Address for correspondence	Enver Roshi Faculty of Public Health, University of Medicine, Tirana, Albania Rr. "Dibres", No. 371, Tirana, Albania E-mail: roshienvi@yahoo.com
Keywords	Analytic surveys, case-control studies, cohort studies, cross-sectional studies, descriptive surveys, experiments, observational studies.
Learning objectives	At the end of the module, students should: • get familiar with the major classifications of epidemiological studies; • get familiar with the major types of epidemiological studies: cohort, case-control and cross-sectional approaches; • know the difference between observational studies and experiments; • know the major epidemiological study designs: applications, advantages and disadvantages.
Abstract	This module deals with the main classifications of epidemiological studies: - Descriptive vs. analytic studies; - Observational studies vs. experiments; - Prospective (forward-looking) design vs. retrospective (backward-looking) approach; - Retrolective vs. prolective data collection. Furthermore, this module outlines the three major types of epidemiological inquiries: • Cohort studies: applications, advantages, disadvantages; • Case-control studies: applications, advantages, disadvantages; • Cross-sectional studies: applications, advantages, disadvantages.
Teaching methods	The teaching method recommended: • the introduction lecture relating to basic definitions and concepts; • distribution of the literature to small group (3-4 students); • the guided discussion within each group and added explanations; • overall discussion.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • training materials: recommended readings, or other related readings; • target audience: bachelor students and master degree students according to Bologna scheme.
Assessment of students	Take-home exercise on the applications, advantages and limitations of the main epidemiological studies.

INTRODUCTION TO EPIDEMIOLOGICAL STUDIES

Enver Roshi, Genc Burazeri, Ervin Toci

Epidemiological studies may deal merely with the distribution of diseases/conditions in human populations (descriptive surveys), and/or with the factors influencing the distribution and the frequency of diseases (analytic surveys: cross-sectional, case-control, cohort, quasi-experiment and experimental studies) (1-3).

Descriptive studies

Descriptive surveys merely portray a situation, e.g. distribution of a disease/condition in a certain population in relation to sex, age, or other characteristics (1,2).

Analytical studies (explanatory studies)

The analytical studies have two features (1-3):

- Tests hypotheses;
- Look for associations based on: a) groups (ecological/correlation studies, trend studies); b) based on individuals (cross-sectional, case-control, cohort, experiments and quasi-experiments).

This is the most classical way of classifying the epidemiological inquiries. However, this classification has little practical value in itself, since the same study could be both descriptive and analytic or, in principle, all studies could be regarded as analytic (e.g. the distribution of a certain condition by sex or age could be regarded as a sort of implicit analysis rather than just description of the observed facts).

The most important way of classifying the epidemiological studies is the one which accounts for the role/control of the researcher over the study. According to this, a major distinction is being made of: a) observational studies, and; b) experiments (1-3).

Observational studies

The observational studies have the following features (2,3):

- The investigator observes the occurrence of the condition/disease in population groups that have assigned themselves to a certain exposure.
- Often most practical and feasible to conduct.
- Carried out in more natural settings – representative of the target population.
- Often, there is little control over the study situation – results are susceptible to distorting influences.

Experimental approach

The experimental studies have the following features (2,3):

- The most powerful study design for testing ethiological hypothesis.
- The investigator exercises control over the allocation of exposure, its associated factors and observation of the outcome.
- For obvious ethical and practical reasons, the possibilities of conducting experiments in human populations are very limited.

Direction of study question

According to the direction of inquiry, epidemiological studies are classified as follows (2):

Prospective (forward-looking) approach – in which (disease free) people who are exposed and non-exposed are followed up and compared with respect to the subsequent development of the disease/outcome under study (2).

Retrospective (backward-looking) approach – in which people with the disease are compared with people without the disease, to determine whether they differ in their past exposure to the (hypothesized) causative factor (2).

Non-directional design – the investigator observes simultaneously the exposure and disease status in the study population (2).

Ambispective design – one primary variable/factor is measured prospectively and the other one retrospectively, or one primary variable/factor is measured both prospectively and retrospectively (2).

Direction of data collection

According to the direction of inquiry, epidemiological studies are classified as follows (2):

Retrolective – data collected before the study design (not necessarily for the purpose of the actual study). This is often the case of routinely collected data in studies involving e.g. occuptational health, or environmental epidemiology studies (2).

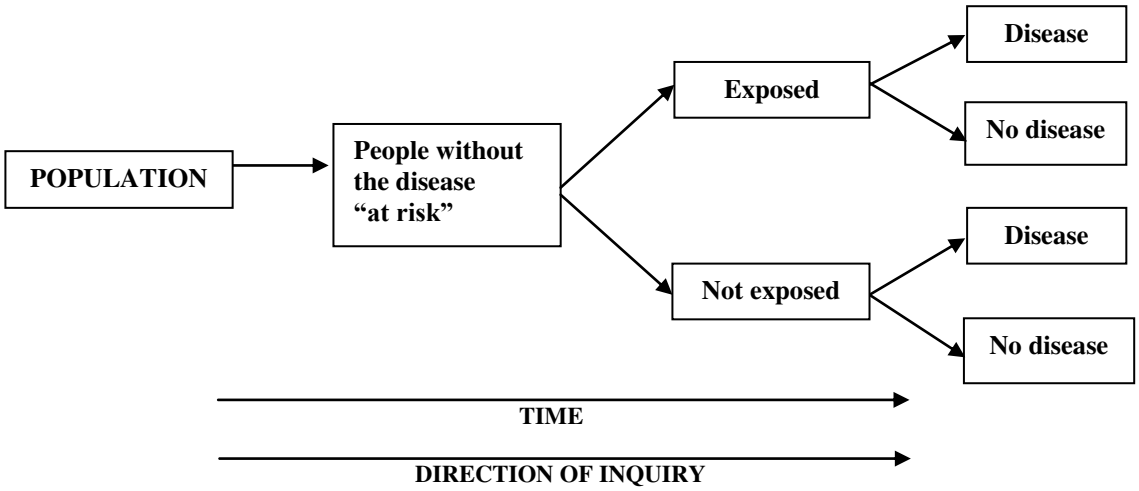
Prolective – data collected after study design (for the purpose of the actual study) (2).

Three major types of epidemiological studies

Cohort studies

The design of a cohort study is sketched below (Figure 1) (4,5):

Figure 1. Design of a cohort study



Characteristics of cohort studies (4-8)

- The essence of cohort studies is the identification of a group of subjects about whom certain exposure information is collected. The group is then followed up in time to ascertain the occurrence of disease/condition of interest.
- For each individual, the prior exposure can be related to subsequent disease experience.
- Since the first requirement of such studies is the identification of the individuals forming the study group – COHORT, prospective or longitudinal studies are usually referred to as cohort studies.

Advantages of cohort studies (4-8)

The main advantages of cohort studies can be summarized as follows:

- Measure incidence and thus permit direct estimation of risk of disease.
- Do not rely on memory for information about exposure status, hence avoid bias due to selective recall.
- Since cohort studies begin with people free of disease, potential bias due to selective survival is eliminated.
- Cohort studies provide a logical approach to studies of causation or effects of treatment.
- Cohort approach can yield information on associations of exposure with several diseases.

Disadvantages of cohort studies (4-8)

Conversely, the main limitations of cohort studies include the following:

- Require large samples to yield the same number of cases that could be studied more efficiently in a case-control study.
- Particularly inefficient for studies of rare diseases.
- Logistically difficult – long follow-up period, often serious attrition to study subjects.
- Direct observation of participants may cause changes in health behavior.
- Possible bias in ascertainment of disease due to changes over time in criteria and methods.
- Very costly.

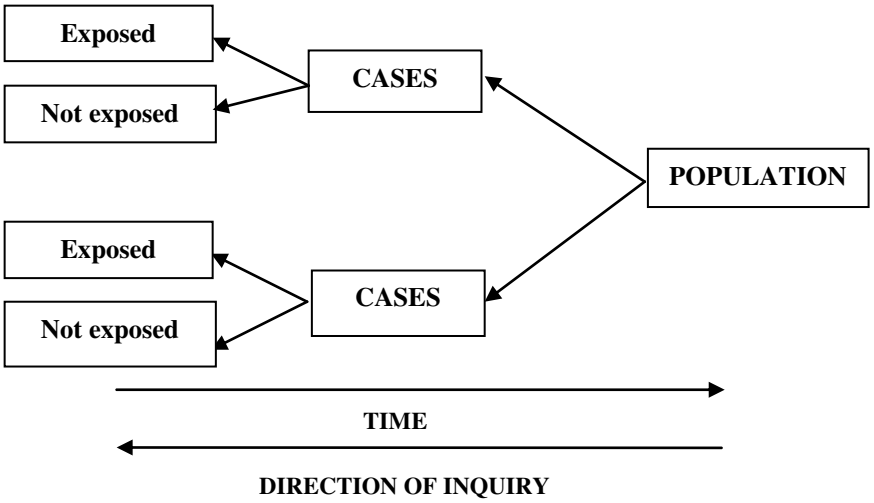
An alternative strategy to the costly and time consuming prospective cohort design is the historical cohort study. A cohort is identified (enumerated) as of some historical point in time and is then followed

over past time to the present. Disease rates and relative risk (RR) can be derived from this type of study as well (4). The retrospective (historical) cohort approach is particularly amenable to the study of exposure-disease relationship for which the exposure group is unusual in some way, e.g. many occupational or environment exposure-disease relationships (4).

Case-control studies

The design of a case-control study is sketched below (figure 2) (9,10):

Figure 2. Design of a case-control study



The case-control study begins with a group of cases of a specific disease. This (the disease) is the starting point of the study, unlike cohort studies where the interest is in drawing a contrast between exposed and non-exposed subjects. Hence, the case-control approach is directed at the prior exposures, which caused the disease and thus proceeds from effect (outcome) to cause (exposure) (9,10).

Advantages of case-control studies

The main advantages of case-control studies are summarized below (9-11):

- Highly informative compared to other designs: several exposures or potential causal agents can be examined.
- Efficient designs (low cost per study) primarily because few subjects are needed to obtain stable estimates of RR.
- Particularly appropriate for studies of rare diseases (e.g. a case-control study with 100 cases of a disease having an annual incidence of 1/1000. A cohort design for this disease would require 1000 persons to be followed up for 100 years or 10000 persons to be followed up for 10 years in order to yield the same number of cases).

Disadvantages of case-control studies

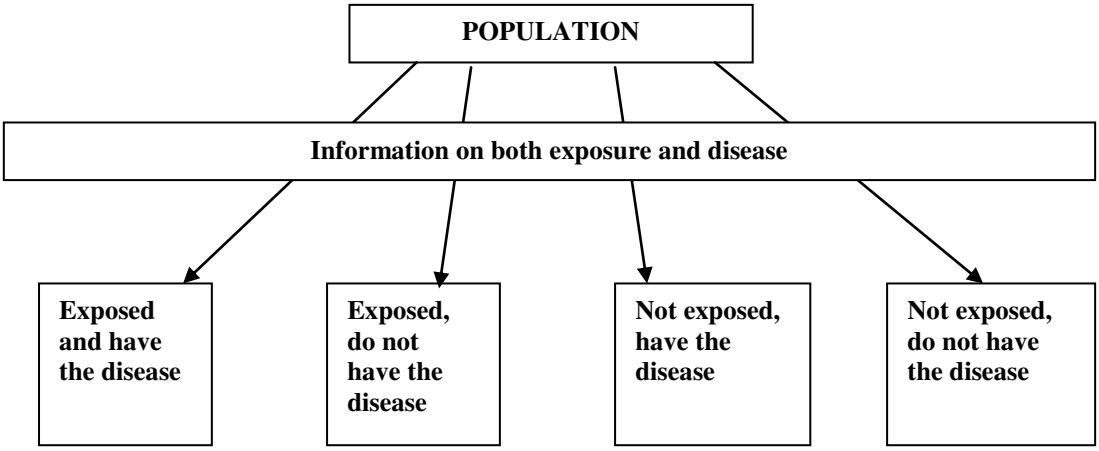
On the other hand, the main limitations of case-control studies are as follows (9-11):

- The absolute frequency of a disease can not be determined. No counts are made of the population at-risk, thus there are no denominators available to obtain the incidence rates. Lacking absolute risks, it is not possible to compare disease rates among different studies, nor is it possible to estimate the attributable risk.
- Particularly subject to bias: selection bias in choosing controls and recall bias (cases may recall better prior exposures than controls).
- "Philosophically" difficult to interpret: the antecedent-consequent relationship (exposure-outcome) is subject to considerable uncertainty.

Cross-sectional studies

The design of a cross-sectional study is sketched below (figure 3) (9):

Figure 3. Design of a cross-sectional study



Cross-sectional design is referred to as non-directional or one point in time survey, where data is collected on both outcome and exposure status of the individuals under study (9). Such studies are useful to describe characteristics of the study population and can generate new ethiological hypothesis. This study design involves disease prevalence. Cross-sectional studies can evaluate the impact of changes in health services during an intervening period. This can be accomplished by conducting a cross-sectional study twice: before and after the intervention (9).

Advantages of cross-sectional studies (9,12)

The main advantages of cross-sectional studies are briefly summarized below

- Describe the distribution of both exposure and outcome in a population (particularly useful for studying frequent outcomes of long duration).
- Provide estimates of the magnitude of a disease problem in a community, which might be very important for the planning of health services.
- Compared with other studies are relatively quick and inexpensive. Often, involve only one-time survey.
- Largely applicable: provision of health care services as well as generation of ethiological hypotheses.

Disadvantages of cross-sectional studies (9,12)

On the contrary, the main limitations of cross-sectional studies include the following:

- Do not measure risk, because this would require incidence data.
- Can not determine cause-effect relationship.
- Current exposure status may be due to changes that have occurred as a result of the disease rather than having led to the disease.
- Diseases of short duration may be missed. Thus, cross-sectional studies are best applied to the study of chronic or persistent conditions.

Exercises

Task 1: Students should be divided up into three groups. Each group should draw a diagram of one of the three major types of epidemiological designs (cohort studies, case-control studies and cross-sectional studies, respectively). Afterwards, in a plennary session, each group should present the design, strengths and limitations of the respective study designs.

Task 2: Each student should search the literature and find out outstanding examples of cohort studies, case-control studies and cross-sectional studies conducted worldwide including their own countries (if applicable).

References and Recommended Readings

1. Last JM, Spashoff RA, Harris SGA. A dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001.
2. Abramson JH, Abramson ZH. Survey methods in community medicine. Churchill Livingstone, New York; 1999.

3. MacMahon B, Trichopoulos D. Epidemiology: principles and methods. 2nd ed. Boston: Little, Brown and Company; 1996.
4. Gordis L. Epidemiology. 4th ed. Philadelphia: Saunders, Elsevier; 2009. Chapter 9: Cohort studies (pp. 167-176).
5. Aschengrau A, Seage GR III. Essentials of epidemiology in public health. 2nd ed. Sudbury: Jones and Bartlett Publishers; 2008. Chapter 8: Cohort studies (pp. 201-228).
6. Munoz A, Nieto FJ. Cohort studies. In: Detels R, Beaglehole R, Lansang MA, Gulliford M, eds. Oxford textbook of public health. 5th ed. Oxford: Oxford University Press; 2009. Pp. 508-525.
7. Szklo M, Javier FN. Epidemiology: Beyond the basics. 2nd ed. Sudbury: Jones and Bartlett Publishers; 2007. Pp. 77-89.
8. Bhopal R. Concepts of epidemiology. 2nd ed. Oxford: Oxford University Press; 2008. Pp. 306-311.
9. Gordis L. Epidemiology. 4th ed. Philadelphia: Saunders, Elsevier; 2009. Chapter 10: Case-control studies and other study designs (pp. 177-200).
10. Aschengrau A, Seage GR III. Essentials of epidemiology in public health. 2nd ed. Sudbury: Jones and Bartlett Publishers; 2008. Chapter 9: Case-control studies (pp. 229-260).
11. Weiss NS. Case-control studies. In: Detels R, Beaglehole R, Lansang MA, Gulliford M, eds. Oxford textbook of public health. 5th ed. Oxford: Oxford University Press; 2009. Pp. 498-507.
12. Kogevinas M, Chatzi L. Cross-sectional studies. In: Detels R, Beaglehole R, Lansang MA, Gulliford M, eds. Oxford textbook of public health. 5th ed. Oxford: Oxford University Press; 2009. Pp. 472-483.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Features of epidemiological studies**
Module: 2.11	ECTS (suggested): 0.15
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Ivan Erzen, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Doncho Donev, MD, PhD, Professor, Director Institute of Social Medicine, Medical Faculty, University of Skopje, Macedonia
Address for correspondence	Lijana Zaletel-Kragelj Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Classifications of epidemiological studies, epidemiological studies, features of epidemiological studies.
Learning objectives	After completing this module students should: • be familiar with the major classifications of epidemiological studies. • understand differences between various types of studies and interpret them.
Abstract	Rarely, prior introducing the readership main study designs, textbooks on epidemiological methods are discussing about different features/characteristics of epidemiological studies. Since one study design has several of them, this could be for the audience sometimes rather confusing. To clarify this aspect of epidemiological studies, we decided to describe these features in more detail.
Teaching methods	An introductory lecture gives the students insight in different features of epidemiological studies and distinction between them. After introductory lectures, students discuss in small groups these features and confront them.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Essay type exam.

FEATURES OF EPIDEMIOLOGICAL STUDIES

Lijana Zaletel-Kragelj, Ivan Erzen, Doncho Donev

Theoretical background

Introduction

There exist a pleiad of textbooks, manuals and other types of books, describing epidemiological concept and its methods in one or another way. Most of them are focused first on different epidemiological measures, and then from one or another perspective on different study designs (ecological studies, cross-sectional studies, case-control studies, cohort studies, clinical trials, community trials, etc.) (1-12). A brief overview on three major types (designs) of epidemiological studies namely cross-sectional, case-control, and cohort studies, is presented also in a previous module of this manual, while some points of view or more detailed description of individual design will be given in the following modules of this book.

Rarely, prior introducing the readership main study designs, these books are discussing about different features/characteristics of epidemiological studies. Since one study design has several of them, this could be for the audience sometimes rather confusing. To clarify this aspect of epidemiological studies, we decided to describe these features in more detail.

Overview of different features of epidemiological studies

Out of numerous different features, two of them could be classified as the most important ones, namely:

- is the study observational or experimental, and
- is the study descriptive or analytical.

But in fact, epidemiological studies have many different features (1-8) and not only the ones just mentioned. These features are used as criteria for classification of epidemiological studies as well. Thus, epidemiological studies could be classified in many different ways. Some of these groups of features are as follows (Figure 1):

1. is the study observational, or experimental (interventional),
2. is the study descriptive, or analytical,
3. is the study cross-sectional, or follow-up,
4. is the study incidence, or prevalence study,
5. is the study question directed into the future, or into the past,
6. is the collection of study data directed into the future, or into the past,
7. is the study using individual, or population/aggregated level data,
8. is the study for generating or testing hypotheses,
9. is the study using permanent data sources, or focused study data,
10. is the study focused on effect, or on exposure,
11. is the study for planning or evaluating an intervention,
12. is the study focused on solving problems at an individual or at a population level.

Besides the first two, classifications listed under the numbers #3-5 could also be met frequently compared to the rest of them. In continuation, we will describe all of them in more detail.

Classifying epidemiological studies by different features

Experimental versus observational studies

Experimental studies are those characterized by assignment of exposure by a researcher. This means that a researcher gains the mastery over the situation (1-8).

Another important characteristic is that experimental studies use the well-known method of randomization for controlling confounding. The importance of randomization is that it leads to a balance of confounders in exposed and non-exposed study groups, providing theoretically unbiased evaluation of exposure-outcome associations. The study and control groups are comparable except in exposure under observation.

Because of their characteristics the experimental studies are the most powerful of all the study designs.

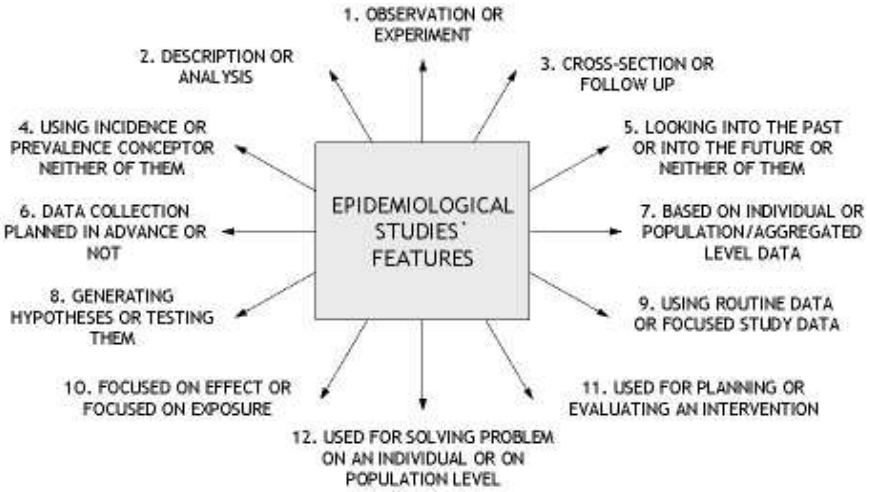
But in spite of their importance, experimental studies have an important disadvantage in human research – they could be extremely unethical and therefore not possible to conduct. Here, observational studies take advantage over experimental studies.

In contrast, observational or non-experimental studies are epidemiological studies that do not involve any intervention, experimental or otherwise. Exposure in this type of studies is not assigned by a researcher. In such a study nature is allowed to take its course, with changes in one characteristic being studied in relation to changes in other characteristics.

Analytic epidemiological methods such as case-control and cohort study designs, are properly called observational epidemiological studies because the investigator is observing without intervention other than to record, classify, count, and statistically analyze results.

An important disadvantage of observational studies is the limited control over the subjects under study and confounding factors that may influence the results substantially. As a consequence, these studies are more susceptible to different types of biases.

Figure 1. Some groups of features of epidemiological studies



Descriptive versus analytical studies

The distinction between "descriptive" and "analytic" studies is one of intent, objective, and approach, rather than one of design. In this respect, data obtained in public health research usually could be explored in a descriptive or analytical mode (8). Data obtained in an analytic study must be first described, and data obtained in a descriptive study can be analyzed to test hypotheses if indicated.

Epidemiological studies designed and concerned primarily to describe the existing distribution of health phenomena in the population, without regard to causal or other hypotheses are usually studies based on routine data. Such studies describe the health conditions and health-related characteristics of populations, typically in terms of person, place, and time. Their results are usually presented in health statistics yearbooks, and similar publications. This information serves as the foundation for studying populations. It provides essential contextual information with which to develop hypotheses, design studies, and interpret results. Surveillance is a particular type of descriptive epidemiology for monitoring change over time.

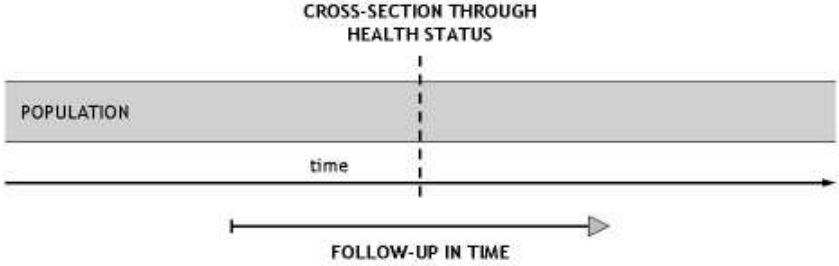
In contrast, analytical studies are usually concerned with testing one or more specific hypotheses, typically whether an exposure is a risk factor for a disease or an intervention is effective in preventing or curing a disease (or any other occurrence or condition of interest) (1-10).

Herewith, we need to emphasize that it is a stereotype that analytic studies (testing hypotheses) take advantage over descriptive studies. As a consequence, descriptive studies are labeled as "less important" than, or "inferior" to analytic studies. However, a well-designed descriptive study should be the first step in a process of investigating health problems of the population, since they are a useful source for generating sound hypotheses for more in-depth and usually more expensive analytical studies.

Cross-sectional or transversal versus follow-up or longitudinal studies

One of the most important features of epidemiological studies is whether the observation is being made as a transversal cross-section through health status of a population, or as a follow-up in time (Figure 2).

Figure 2. Schematic presentation of cross-sectional and follow-up studies



Cross-sectional or transversal studies are those studies that examine the relationship between diseases or other health-related states at a particular time, being a moment or a period (1). In this design subjects are sampled without respect to disease status and are studied at a particular point in time (8), and the presence or absence of an outcome, as well as presence or absence of exposure is observed in the same point in time.

The term "cross-sectional study" usually refers to studies at the individual data level, even though ecologic studies at aggregate level are typically (though not necessarily) cross-sectional, as well.

The target population is usually one whose identity is of some wider interest (e.g., a political or geographical entity, a profession or workforce, or a major organization, but may not necessarily be so) (8).

In cross-sectional studies, the current status of individuals could be examined in relation to some current or past exposure. When these studies are used with an analytical purpose, one should be cautious when interpreting the relationship between outcomes and exposures, especially the causal ones, since temporal sequence of cause and effect cannot necessarily be determined by this study design (1,4).

These studies are most useful for conditions that are not rapidly fatal, not terribly rare, and/or not routinely brought to medical attention (e.g., elevated blood pressure, elevated blood cholesterol, etc.) (8).

In contrast, in follow-up or longitudinal studies people without the disease at the beginning of the observation time (usually referred to as "at risk") are followed-up to observe development of the disease over time. If the population followed is a defined group of people (a "cohort"), then the study is referred to as a cohort study.

Special type of longitudinal studies are so-called ecological longitudinal studies (3) that are studies made on ongoing frequent cross-sectional studies (surveillance or monitoring) to measure trends in disease rates over many years in a defined population. By comparing the trends in disease rates over time and considering other changes in the population, it could be determined the impact of these changes on the disease rates.

Prevalence versus incidence studies

The term "prevalence studies" is referring to cross-sectional studies at the individual data level. The frequency of an outcome variable in this type of studies is measured in terms of prevalence. Prevalence is a common term for a group of measures which are quantifying the situation (state) of a given health phenomenon (e.g. a disease, a disorder, an unhealthy health behaviour etc.) at a designated time (at a specified moment, or at any time during a specified period). The detailed description of prevalence measures (e.g. prevalence risk, prevalence rate, prevalence odds), as well as that of characteristics of cross-sectional studies at the individual data level, are beyond the scope of this module, and are given in other modules of this book.

Similarly, the term "incidence studies" is referring to studies, also known as follow-up, longitudinal, or cohort studies. The frequency of an outcome variable in this type of studies is measured in terms of incidence. Incidence is a common term for a group of measures which are quantifying a breakout of new cases of a health phenomenon (e.g. a disease) under observation during a specified period in a specified group of persons. The detailed description of incidence measures (e.g. incidence risk, incidence rate, incidence odds, and incidence density), and characteristics of cohort studies, are beyond the scope of this module. They are provided in other modules of this book.

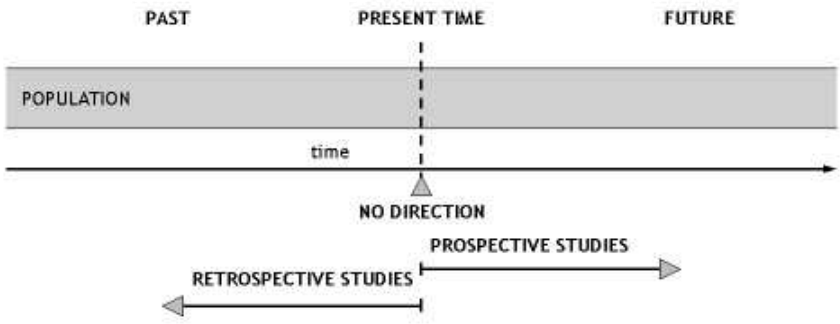
In contrast, case-control studies are not incidence or prevalence studies, and are intended to observe past exposure.

Retrospective versus prospective, versus non-directional studies

One very important feature concerns the timing of collection of exposure information. In this respect, we distinguish between retrospective, prospective, and non-directional studies (Figure 3).

A retrospective study is a study that looks backwards in time and examines exposures to suspected risk or protection factors in relation to an outcome that is established at the start of the study. Thus, in a retrospective study design, the outcome of interest has already occurred at the time the study was initiated. Hence, we find people that have a disease under observation and try to figure out why they got the disease.

Figure 3. Schematic presentation of retrospective, prospective, and non-directional studies



An investigator conducting a retrospective study typically utilizes administrative databases, medical records, or interviews with patients who are already known to have a disease or condition.

The biggest problem in a retrospective study is that some of the information that we need may be hard to get, or it is subject to the so-called "recall bias". We have to rely on patients to recall things that may have happened many years ago. In contrast, a prospective study looks forward in time. In this study design, we select a group of subjects without a condition under observation and observe them over a specified period if they develop the condition after they were exposed to a suspected risk or protective factor.

One of the disadvantages of this study design is that in a case the outcome under observation has a long pre-clinical phase it could take a long time to accumulate sufficient data to get correct and strong conclusions. When we are studying a disease that takes a long time to appear, we usually need to use a retrospective study, and not a prospective one.

The outcome of interest also should be common; otherwise, the number of outcomes observed will be too small to be statistically meaningful (indistinguishable from those that may have arisen by chance).

All efforts should be made to avoid sources of bias such as the loss of individuals to follow up during the study. Prospective studies usually have fewer potential sources of bias and confounding than retrospective studies.

Despite these disadvantages, prospective study designs are the best design for establishing relationships between outcomes of interest and exposure variables.

The third study design in this group is non-directional, in which outcome(s) and exposure(s) are observed at the same time. Transversal or cross-sectional studies are non-directional studies. We have already discussed some characteristics of this type of studies, while some further details will be discussed in a separate module in this book.

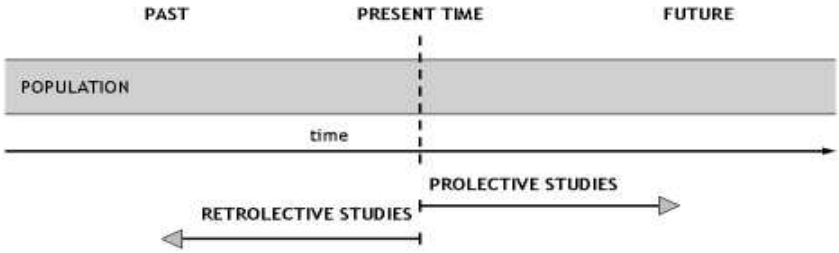
Retrolective versus prolective studies

A less frequently used feature and corresponding classification relates to the mode of gathering of data. In this respect, we distinguish between retrolective and prolective studies (1) (Figure 4).

According to Feinstein, who coined both terms at the beginning of 1980s (1), this classification describes more precisely the actions of researchers than more commonly used terms "retrospective" and "prospective" studies.

The term "retrolective studies" relates to data gathered from medical records or other sources when data collection took place without prior planning for the needs of a present study, while the term "prolective studies" relates to data collected by planning in advance.

Figure 4. Schematic presentation of retrolective and prolective studies



Studies based on individual level data versus studies based on population/aggregated level data

Another feature and corresponding way of classification of epidemiological studies involves the level of measurement. According to the level of measurement of variables that enter the studies, studies are classified as (11-21):

- individual data level studies;
- aggregated data level studies: measures in these studies are summaries of attributes calculated from data on individuals for whole populations, usually in well-defined geographic or administrative regions (e.g. countries, communities etc.). Examples for that kind of measurements would be: mean income, percentage of families below the poverty line or mean number of household members;
- group data level studies: measures in these studies are estimates of (environmental) attributes that have individual analogues. Usually these measures are obtained from different environmental surveys. Examples for that kind of measurements would be: maximum daily exposure to ozone, mean annual exposures to radon gas; daily mean levels of environmental tobacco smoke in public buildings;
- population level studies: measures in these studies are attributes that pertain to groups and do not have analogues at the individual level. Examples for that kind of measurements would be: total area of green space; number of private medical clinics; population density.

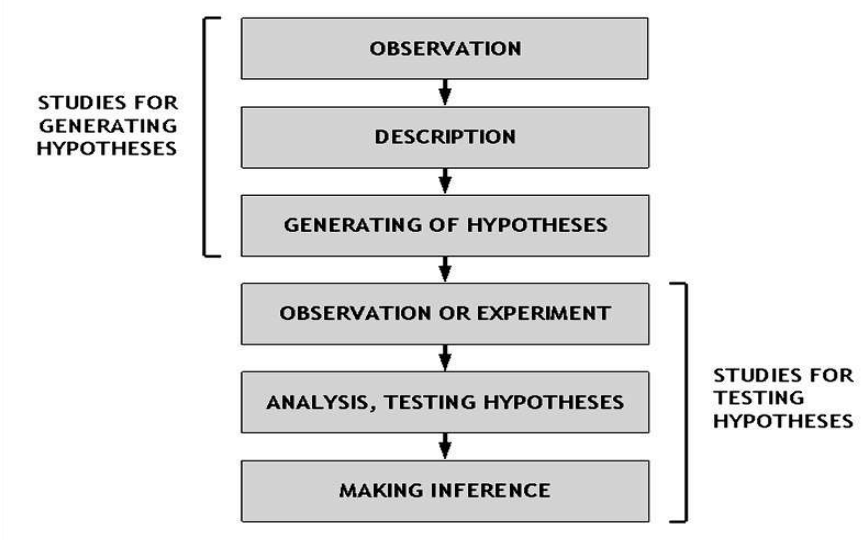
Studies for generating hypotheses versus studies for testing hypotheses

This classification of epidemiological studies is less frequently used, since studies for testing hypotheses take advantage over studies for merely generating hypotheses.

Usually, epidemiological studies for generating hypotheses are basically descriptive, and as such are kept in the background in comparison to analytical studies. This is from historical point of view understandable to a certain extent, since public health is in many respects tightly connected to biomedicine, where studies for testing hypotheses are taking huge advantage, especially the experimental ones.

In spite of this fact, the importance of studies for generating hypotheses lies in the fact that they provide essential contextual information for sound analytical studies for testing hypotheses which are mostly also more expensive (Figure 5).

Figure 5. Schematic presentation of position of studies for generating hypotheses, and studies for testing hypotheses in epidemiologic inference



One should be aware that both types of epidemiological studies are part of the same process which could be more sound and effective when both types of studies complement each other.

Among others, very important studies for hypotheses generation are routine analyses of vital statistics and other notifiable events, periodic surveys of health status, knowledge, beliefs, attitudes, practices, behaviours, environmental exposures, and health care services performance, as well as ecological studies that compare information across geographical or administrative units.

In contrast, epidemiological studies for testing hypotheses are basically analytical. The most respected are experimental and cohort studies. Hypotheses could be tested also in case-control, and cross-sectional studies.

This distinction is more and more important, especially in the context of evidence-based public health. Formulation of sound hypotheses on the basis of available data, e.g. on surveillance and monitoring data that are later verified by in-depth analyses is becoming an imperative.

Studies based on routine data versus studies based on data focused on analyzing a specific problem

A lot of data in public health is gathered in the frame of routine surveillance of major public health events. Basically, routine data are meant to describe major health conditions and health-related characteristics of populations, typically in terms of person, place, and time. As stated previously, their results are usually presented in health statistics yearbooks, and similar publications. Nevertheless, they could be logically and reasonably used also for other more in-depth analyses.

On the other hand, for less important or evolving problems, specific studies are more appropriate, since they are less expensive.

Studies focused on exposure rather than on effect

In the past, environmental epidemiology and occupationally epidemiology have been mainly oriented in studying associations between disease and environmental agents. A broader approach is currently envisaged, which is primarily focused on exposure circumstances and which considers as dependent variables all possible health effects of environmental agents to which populations are exposed (22,23).

There are several reasons for the shift from disease- to exposure-centred environmental epidemiology:

1. Firstly, particularly in developed countries, degenerative, chronic diseases (such as cancer, lung emphysema, etc.) have become the prevailing pathology: the aetiology of many of these conditions is multifactorial, i.e. no specific hazard can be considered as a necessary cause. To further complicate the picture, many environmental hazards (e.g. excess dietary fat, asbestos, etc.) are causally associated with more than one disease.
2. Secondly, most environment-induced conditions are dose-related. For a given hazard, there may well be exposures either low enough, or of short enough duration, as to be negligible in terms of risk. It has also become obvious that ill-effects are frequently the result of interaction (addition, synergism, antagonism, etc.) between different hazards. For the same exposure to a given hazard, the risk may differ according to whether other hazards are present or not.

3. Thirdly, analytical techniques for measuring pollutants in the environment have been used more and more, and their sensitivity has increased by several orders of magnitude. Consequently, there has been a dramatic increase in hazard-specific environmental data requiring risk evaluation.
4. Finally, health authorities, public opinion, and the scientific community have become increasingly concerned by the number of environmental contaminants for which potentially deleterious effects are unknown or poorly understood.

Studies for planning versus studies for evaluating public health interventions

Again, the distinction between "studies for planning public health interventions" and "studies for evaluating public health interventions" is one of intent, objective, and approach, rather than one of design. In this respect, data obtained in public health research usually could be explored for planning or evaluation.

Clinical epidemiologic studies versus public health epidemiologic studies

Clinical epidemiological studies are studies conducted in clinical settings, usually by clinicians, with patients as the subjects of the study. They apply epidemiological principles and methods to problems observed in clinical medicine. Their intention is to use the information from classical epidemiology to aid clinical decision making (1). In contrast, classic (public health) epidemiological studies are intended to identify causes of diseases, and measure risk (1).

Some other features of epidemiologic studies

Along with the features of epidemiologic studies presented so far, there exist other features as well. It could be worth to mention at least two of them, being pragmatic and explanatory study:

- according to Last et al. (1), pragmatic study is a study aimed at providing a basis for decisions about health care, or evaluating previous action (interventions);
- according to Last et al. (1), explanatory study is a study aimed at explaining rather than merely describing the situation of a certain health problem by isolating the effects of specific variables and understanding the mechanisms of action.

Exercises

Task 1

Students carefully read the theoretical background of this module and the recommended readings.

Task 2

Students make groups for crossover methods of discussion. Every student is labelled with two tags including letters and figures: A1, A2, A3, B1, B2, B3, C1, C2, C3, etc.

For the first part of this task, all As, Bs, and Cs work together. The first part of the task is:

- GROUP A: discuss differences of experimental versus observational studies in relation to descriptive versus analytical studies;
- GROUP B: discuss differences of cross-sectional or transversal versus follow-up or longitudinal studies in relation to prevalence versus incidence studies;
- GROUP C: discuss differences of retrospective versus prospective versus non-directional studies in relation to retrolective versus prolective studies.

For the second part of this task, all 1s, 2s, and 3s work together. The second part of the task is the same as the first, but groups are different:

- GROUP 1: discuss differences of experimental versus observational studies in relation to descriptive versus analytical studies;
- GROUP 2: discuss differences of cross-sectional or transversal versus follow-up or longitudinal studies in relation to prevalence versus incidence studies;
- GROUP 3: discuss differences of retrospective versus prospective versus non-directional studies in relation to retrolective versus prolective studies.

Task 3

In large group discuss other things which deserve/warrant classifications.

References

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 4, 2013.
3. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia, W.B. Saunders company; 2001.
4. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.

5. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
6. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization; 1998.
7. Hennekens CH, Buring JE. Epidemiology in medicine. Boston, Little, Brown and Company; 1987.
8. Schoenbach, VJ, Rosamond WD. Understanding the fundamentals of epidemiology. An evolving text. Chapel Hill: University of North Carolina, School of Public Health, Department of Epidemiology, 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 4, 2013.
9. Marusić M, Petrovecki M, Petrak J, Marusić A. Introduction to the scientific research in medicine (in Croatian). Zagreb: Medicinska naklada; 2000.
10. Gjorgov A, Gjorgov N. Research methodology and clinical epidemiology. Skopje: Grigor Prlicev Publishing House; 1999.
11. McMichael AJ. Setting environmental exposure standards: The role of the epidemiologist. Int. J Epidemiol 1989;18:10-16.
12. Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. Environ Health Perspect 1993; 110:23-38.
13. Armstrong BK, White E, Saracci R. Principles of Exposure Measurement in Epidemiology. Monographs in Epidemiology and Biostatistics, New York Oxford University Press; 1994.
14. Susser M. The logic in ecological: I. The logic of analysis. Am J Public Health. 1994;84:825-9.
15. Susser M. The logic in ecological: II. The logic of design. Am J Public Health. 1994;84:830-5.
16. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health. 1995;16:61-81.
17. Susser M, Susser E. Choosing a future for epidemiology: I. Eras and paradigms. Am J Public Health. 1996;86:668-73.
18. Susser M, Susser E. Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology. Am J Public Health. 1996;86:674-7.
19. Pearce N. Traditional epidemiology, modern epidemiology, and public health. Am J Public Health. 1996;86:678-83.
20. Künzli N, Tager IB. The semi-individual study in air pollution epidemiology: a valid design as compared to ecologic studies. Environ Health Perspect. 1997;105:1078-83.
21. Ribeiro MC, Pereira MJ, Soares A, Branquinho C, Augusto S, Llop E, et al. A study protocol to evaluate the relationship between outdoor air pollution and pregnancy outcomes. BMC Public Health. 2010;10:613.
22. Morgenstern H. Ecologic studies. In: Rothman KJ, Greenland S, editors. Modern Epidemiology. Philadelphia: Lippincott-Raven; 1998. p.459-80.
23. Harvey T, Mahaffey KR, Velazquez S. Dourson M. Holistic risk assessment: an emerging process for environmental decisions. Regul Toxicol Pharmacol 1995;22:110-7.

Recommended readings

1. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization; 1998.
2. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 4, 2013.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia, W.B. Saunders Company; 2001.
5. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
6. Schoenbach, VJ, Rosamond WD. Understanding the fundamentals of epidemiology. An evolving text. Chapel Hill: University of North Carolina, School of Public Health, Department of Epidemiology, 2000. Available at URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 4, 2013.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Ecological studies**
Module: 2.12	ECTS (suggested): 0.1
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Ivan Erzen, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Lijana Zaletel-Kragelj Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Ecological fallacy, ecological study, epidemiologic study.
Learning objectives	After completing this module students should: • know the definition and characteristics of epidemiological studies; • be familiar with strengths and weaknesses of epidemiological studies; • know how to get data for performing epidemiological studies, • know how to prepare data for an epidemiological study, and • know how to interpret results of an epidemiological study.
Abstract	Ecological study is an epidemiological study in which the units of analysis are populations or groups of people, rather than individuals. They are very applicable in situations in which data are not available at an individual level, but one should be careful when interpreting results of this type of epidemiological studies. The module is describing strengths and limitations of epidemiological studies, as well as procedures for analyzing data in the SPSS programme.
Teaching methods	An introductory lecture gives the students a first insight into the characteristics of ecological studies. The theoretical knowledge is illustrated by a case study. After introductory lectures, students first carefully read the recommended readings. Afterwards, they discuss the characteristics of ecological studies with other students. In continuation, they perform an ecological study analysis by themselves.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet, statistical programme and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire, and output of analysis of an ecological study with its interpretation.

ECOLOGICAL STUDIES

Lijana Zaletel-Kragelj, Ivan Erzen

Theoretical background

Introduction

When we meet the term "ecological study" for the first time without knowing basic principles of epidemiology, the very first association is on ecology, which is defined as the study of the relationship among living organisms and their environment (1). A branch of ecology is human ecology, which is the study of human groups as influenced by environmental factors, including social and behavioural factors. But this association is not completely accurate. "Ecological studies" as a type of epidemiological studies are not related only to studying the influence of natural environment (mostly physical component on people). As a special type of epidemiological study, ecological studies are useful in analysis in various situations not only in the one just described.

It would be logical and reasonable to think about this term in a broader, and in a narrower sense. Meant in their broad meaning, ecological studies are studies of influence of environmental factors on human health, while in a narrow, technical meaning, ecological studies are a special type of epidemiological studies. In this module, the narrowest meaning will be discussed. However, this module describes only the most simple view on the ecological studies. A more detailed view is far too complex to be presented here. Those who wish to deepen their knowledge will find relevant literature in the recommended readings list.

About ecological studies

Definition and description

According to Last et al. (1), ecological study is a study in which the units of analysis are populations or groups of people, rather than individuals. A similar definition is given by Bailey et al. (2): ecological studies are observational epidemiological studies that consider the characteristics of a disease and risk factors measured at the population rather than the individual level. They could be descriptive or analytical.

Thus, in the case of this type of epidemiological studies, if confronted to other types of epidemiological studies, the group of individuals, and not the individual person, is the unit of observation and analysis. In this case, individual measurements are aggregated. Afterwards, an aggregated measure is used in ecological studies (e.g. average or median value of individual values, or percentage of people with observed state). However, this is not the only type of data that can be used in the ecological studies.

Aggregation is usually carried out in a geographical region, or administrative region, as well as in different types of settings, e.g. health care settings, schools, etc.

Ecological studies have been conducted by social scientists for more than one century, and have been used extensively by epidemiologists in many research areas. Nevertheless, the distinction between individual level and group/population level (ecological) studies and the inferential implications are far more complicated and subtle than they first appear (3).

Characteristics

Level of measurement of entry data

According to the level of measurement of variables that enter the ecological studies, there are three different types of measures (4-6):

- aggregated level measures: summaries of attributes calculated from data on individuals for whole populations, usually in well-defined geographic or administrative regions (e.g. countries, communities etc.). Examples for that kind of measurements would be: mean income; percentage of families below the poverty line or mean number of household members;
- group level measures: estimates of (environmental) attributes that have individual analogues. Usually these measures are obtained from different environmental surveys. Examples for that kind of measurements would be: maximum daily exposure to ozone, mean annual exposures to radon gas; daily mean levels of environmental tobacco smoke in public buildings;
- population level measures (measures: contextual): attributes that pertain to groups and do not have analogues at the individual level. Examples for that kind of measurements would be: total area of green space; number of private medical clinics; population density.

Studies may include also variables of different levels. The outcome variable in an ecological study could be measured on a quantitative scale (e.g. percentages, epidemiological rates), or qualitative (2,7).

Ecological studies purposes

The main purposes of using a study at a population level are as follows (2,7-10):

- to study data that could be obtained only at a group-level; health-related data are sometimes available only at the group level (e.g. water or air pollution, percent of green areas in the community etc.);

- to study group-specific effects; this is important since the public health interventions are performed at the group level rather at the individual level;
- to assess very roughly a negative phenomenon which is perceived at a community level, and generate hypotheses for further investigation;
- to investigate differences between populations – in some health phenomena differences are greater between populations than within them (e.g. due to differences in culture or health care system);
- to describe patterns or trends on a geographic or administrative level;
- to explore potential associations between community-level risk factors and disease.

Data sources for ecological studies

Data for ecological studies are obtained from (2,7-9):

- most frequently ecological studies are performed on routine data, since the valuable information about disease and exposure could often be abstracted from published statistics on international, national, regional, or local level. This means, that ecological studies usually do not require expensive or time consuming data collection. These routine data could be first obtained on individuals and than aggregated, or are by nature data that are measured in natural environment, i.e. measurements of air pollution in vicinity of industry;
- data could also be obtained from periodical surveys, like health interview surveys about health behaviour.

Data could be obtained from one source only, or by combining different sources, and could be collected at different times for different purposes.

Advantages and disadvantages

Like other epidemiological studies, ecological studies have some advantages and some disadvantages (2,7-10), which are presented in Table 1.

Table 1. Selected advantages and disadvantages of ecological studies

ADVANTAGES	DISADVANTAGES
1. quick and relatively inexpensive; 2. may be able to use readily available data; 3. useful in hypothesis generation; 4. allow estimation of effects not easily measurable for individuals; 5. permit exploratory analyses of potential factors in disease etiology; 6. appropriate when inferences are to be made about groups and not individuals; 7. useful for social scientists as well as epidemiologists; 8. useful in evaluation of new policies.	1. not able to analyze information on important factors that may be associated with the observed outcome because data are already collected for other purposes, thus difficult to control for confounders; 2. do not provide information about the relationship between risk factor levels and disease in individuals; 3. presence of so called "ecological fallacy" – association observed between variables on an aggregate/population level does not necessarily represent the association at an individual level; 4. exposures and outcomes are not measured on the same individuals; 5. in longitudinal ecological studies migration patterns over time could influence (e.g. diminish) the difference between observed groups.

Methods of analysis of ecological studies

Graphical presentation

Essential part of ecological study analysis is a graphical presentation. In fact, whenever observing the relationship between two quantitative variables, the first step is to plot a diagram. In the case of relating two variables, the diagram/chart is the joint distribution two-dimensional graph, called "scattergram", "scatter diagram", or "scatter plot" (9,11-13). The chart establishes the relationship of a dependent variable to an independent variable. The dependent variable is plotted on the vertical y-axis; the independent is plotted on the horizontal x-axis. Each point represents a place where the dependent and independent variables intersect. In Figure 1a, some general examples of scatter plots are presented.

The dispersion (the scatter) of points of intersection of variable X and variable Y could express a pattern that could be summarized by a mathematical model. In the case presented in Figure 1, a straight line is the proper mathematical model (Figure 2b), the so-called "regression line". The equation of this mathematical model on the sample level is as follows (Equation 1):

$$y = a + bx \qquad \textbf{Equation 1.}$$

The presented relationship is only the most simple, being linear. There exist several others, but this issue is beyond the scope of this module.

A typical scatter plot in ecological studies has intersection points labelled. An example will be presented in case studies.

Correlation

Regarding the nature of the variables that enter the ecological studies, it is logical to use as an analytical method, the statistical method called "correlation" (7,9,11-13), which measures the strengths of association between two variables, or in other words the grade of dispersion of intersection points around the mathematical model. The outcome measure is referred to as "correlation coefficient" labelled at sample level as "r", if it is calculated using a parametrical method (Pearson's correlation coefficient). The equation is as follows (Equation 2):

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$ **Equation 2.**

The value of Pearson's correlation coefficient lies between 0 and 1. The value 0 indicates that there is no linear relationship between variables, while the value 1 indicates the strongest relationship. In this case all intersection points lie on the regression line. In Figure 1a, the value of correlation coefficient is near 0, while in Figure 1c is rather close to 1.

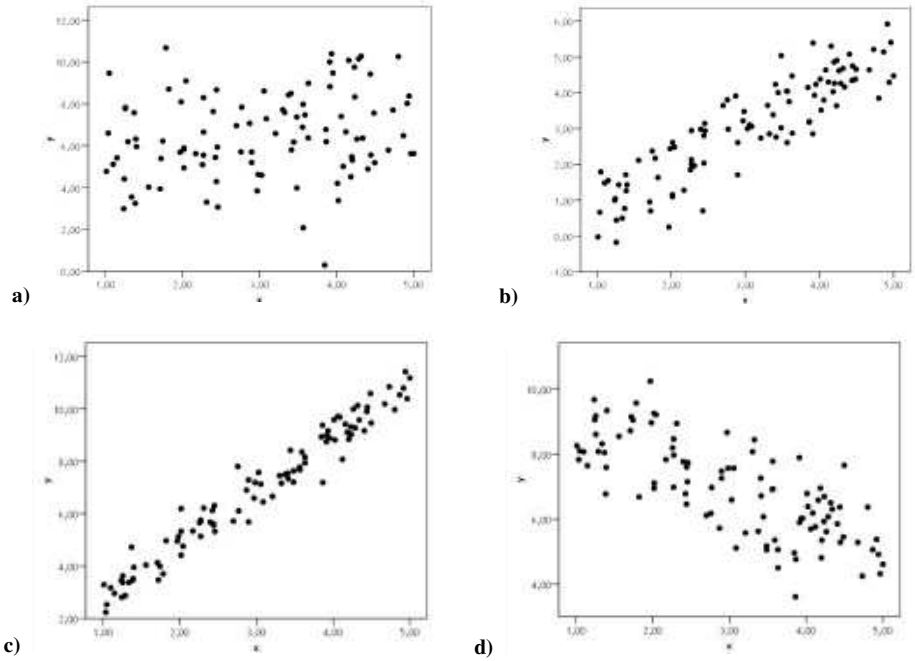
The value of Pearson's correlation coefficient could have positive or negative signs. If with increasing of values of variable X values of variable Y are increasing (Figure 1b, 1c, 2a, 2b), the sign is positive, otherwise it is negative (Figure 1d).

Pearson's correlation coefficient is only one type of correlation coefficient. The detailed description of characteristics of various types of correlation analysis is beyond the scope of this module. It is assumed that students are familiar with the basic statistical methods, including correlation.

In the case of ecological studies the correlation is a special one, called "ecological correlation". According to Last et al. (1), ecological correlation is a correlation in which the units studied are populations rather than individuals. Correlations found in this manner may not hold true for the individual members of these populations.

Because the method of analysis is correlation, ecological studies are frequently called also "correlational studies" (8).

Figure 1. An example of a scatter plot with various degree of dispersion of intersection points (Figures 1a - c). In Figures 1a - c, the relationship between variables X and Y is positive, while in Figure 1d it is negative



Interpretation of ecological studies

Although ecological studies are easily and inexpensively conducted, the results are often difficult to interpret.

In interpretation of results of epidemiologic studies one should be extremely cautious and careful. In fact, ecological studies may be a useful pointer to further research, but conclusions derived from them must be

interpreted wisely. Primarily, we need to be aware that research question in an ecological study is about a population, and not about an individual (Example 1).

> Does the overall occurrence of disease X in a population correlates with occurrence of the exposure in the population? **Example 1.**

It should be pointed out that an ecological study design does not enable to draw any conclusions on the etiological factor of the observed phenomena, otherwise, there is a risk of the so-called ecological fallacy.

Ecological fallacy

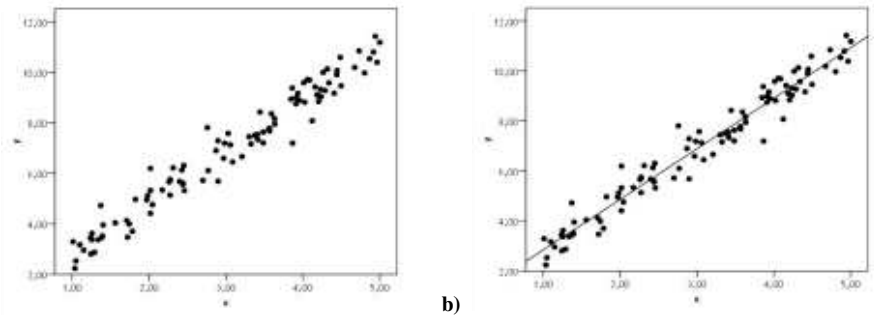
In ecological analysis, errors of inference may result because associations may be artifactually created or masked by the aggregation process.

The ecological fallacy, also known as aggregation bias or ecological bias (1), is the mistaken assumption that a statistical association observed between two ecologic (group-level) variables is equal to the association between the corresponding variables at the individual level. This assumption is often made implicitly or explicitly when using ecologic data to make inferences about the biologic (individual-level) effect of an exposure on the risk of a disease or other health outcome (1,2,7-9).

In extreme situation, an association at one level may disappear at another, or even be reversed. Suppose, for example, we observe a positive ecologic association between exposure prevalence and the rate of a disease across many regions (groups). The magnitude and direction of the association between exposure status and disease risk within regions (at the individual level) could be different from the ecologic association, even if there is no error in measuring either ecologic variable. Just because the disease rate is higher in regions with a larger exposure prevalence does not mean that exposed individuals are at greater risk of disease than are unexposed individuals. It is possible that the risk is particularly high for unexposed individuals living in regions with relatively high exposure prevalence. The underlying problem of the ecologic fallacy, therefore, is that each group is not entirely homogeneous with respect to the exposure. If every region were made up entirely of exposed individuals or unexposed individuals, then there would be no ecologic fallacy because information on the joint distribution of exposure and disease within groups would not be missing.

In conclusion, the aggregation of data that defines ecological studies results in an information loss that can lead to ecological bias or ecological fallacy, respectively. It is due to the inability of ecological data to characterize within-area variability in outcomes, exposures and confounders, when available. The only way to overcome such bias, while avoiding assumptions, concerning the missing information, that could not be checked, is to regard ecological studies as first step in analyzing the problem, and try to supplement the ecological-level information with data at individual-level.

Figure 2. In Figure 2b, a mathematical model (regression line) is added to a scatter plot from Figure 2a



a) **b)**

Meaningless correlations

In ecological studies, meaningless correlations can occur. This kind of correlations sometimes occur when social, economic, or technological changes have the same trend over time as incidence or mortality rates.

Other types of ecological studies

Usually, under the term “ecological study” we understand the ecological studies which are described in this module. However, there exist several types of ecological studies (14). The design just described is in fact the geographical ecological study design (the units of observation at the population level are geographical or administrative units). Another type is temporal ecological study design (the units of observation at the population level are time units, i.e. days). By these studies we can observe trends in disease rates over many years in a defined population. The detailed description is beyond the scope of this module. Both case studies presented in this module are examples of geographical ecological studies.

Rationale for conducting ecological studies, and purpose of this type of studies in public health

There are several reasons for the widespread use of ecologic studies in epidemiology, despite frequent cautions about their methodological limitations (2,3,7-10):

1. Low cost and convenience:
   Ecologic studies are inexpensive and take little time because various secondary data sources, each involving different type of information needed for the analysis, can easily be linked at the aggregate/population level. For example, data obtained from population registries, vital statistics records, large sample surveys, and the census are often linked at the state, county or census-tract level.
2. Measurement limitations of individual-level studies:
   In environmental epidemiology and other research areas, we often cannot accurately measure relevant exposures of doses at the individual level for large numbers of subjects – at least not with available time and resources. Thus, the only practical way to measure the exposure may be ecologically. This advantage is especially true when investigating apparent clusters of disease in small areas (15). Sometimes, individual-level exposures, such as dietary factors, cannot be measured accurately because of substantial within-person variability; yet ecologic measures might accurately reflect group averages.
3. Design limitations of individual-level studies:
   Individual-level studies may not be practical for estimating exposure effects if the exposure varies little within the study area. Ecologic studies covering a much wider area, however, might be able to achieve substantial variation in mean exposure across groups.
4. Interest in ecologic effects and hypotheses generation:
   As noted above, the stated purpose of a study may be to assess an ecologic effect; i.e. the target level of inference may be ecologic rather than biologic – to understand differences in disease rates among populations. Ecologic effects are particularly relevant when evaluating the impacts of social processes or population interventions such as new programs, policies, or legislation. However, an interest in ecologic effects does not necessarily obviate the need for individual-level data.
5. Study group-specific effects:
   This is important since public health interventions are performed at the group level rather at the individual level.
6. Simplicity of analysis and presentation:
   In large complex studies conducted at the individual level, it may be conceptually and statistically simpler to perform ecological analyses and to present ecological results than to do individual level analyses. For example, data from large periodic surveys are often analyzed ecologically by treating some combination of year, region, and demographic group as the unit of analysis.

Despite several practical advantages of ecologic studies, there are many methodological problems that severely limit causal inference, including ecologic and cross-level bias, problems of confounder control, within-group misclassification, and lack of adequate data, temporal ambiguity, co-linearity, and migration across groups (16).

Case studies

Case study 1: Standardized death rate in relation to GDP in countries of WHO, Eruopean Region

Introduction

As an example of an ecological study could be observation of the relationship between different measures/indicators, available on the level of countries, i.e. the indicators, available from Health for All Database of European Region of the World Health Organization (17). For this module, we have chosen to present the relationship between standardized death rate (SDR) from all causes in all age groups per 100,000 population, and Gross domestic product (GDP) per capita (in US$), in countries of European Region of World Health Organization. The data for the year 2001 were chosen since for this year they were available for most of the countries of the region.

Scatter plot

In Figure 3, the scatter plot of SDR from all causes in all age groups per 100,000 population, and GDP per capita, in countries of European Region of World Health Organization, is presented.

From this scatter plot can be observed that the relationship between the variables presented is not linear, so further analysis will not be simple, and the Pearson's correlation coefficient that assumes linearity of the relationship cannot be used.

Correlation analysis

One method to overcome the problem of non-linearity is that countries are classified in three groups according to GDP. It would be reasonable to make following groups:

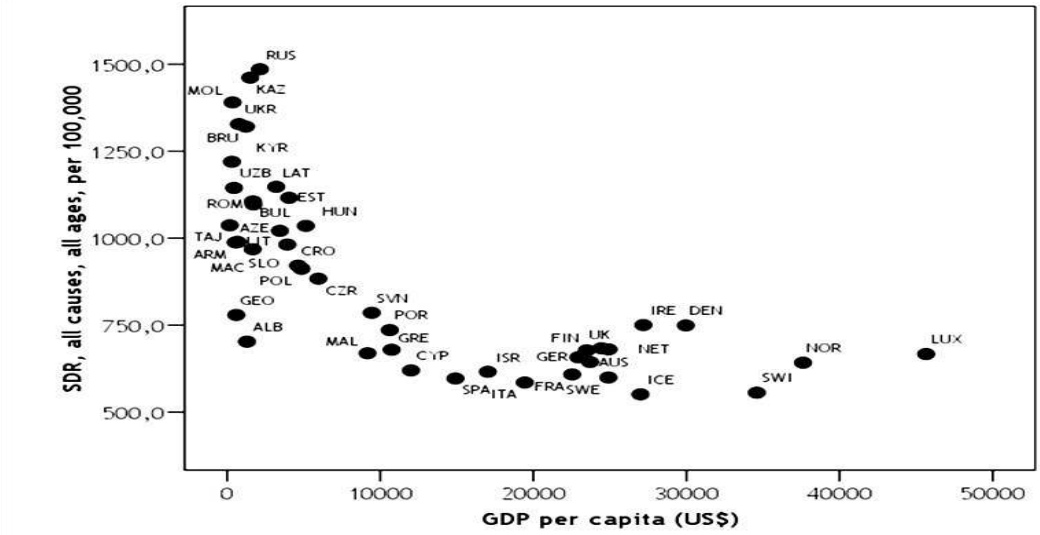
- GDP up to 2999 US$ per capita;
- GDP 3000-7999 US$ per capita, and

- GDP 8000 US$ per capita or higher.

The boundaries were set arbitrary and only for the purposes of demonstration of one way of analysing ecological study data, and do not in any case mean that these boundaries are argumented.

In Table 2, values of correlation coefficients (r) in these groups are presented.

Figure 3. Standardized death rate (SDR) from all causes in all age groups per 100,000 population in relation to Gross domestic product (GDP) (in US$), in countries of WHO, European Region (data for 2001)



LEGEND: ALB - Albania, AND - Andorra, ARM - Armenia, AUS - Austria, AZE - Azerbaijan, BRU - Belarus, BEL - Belgium, BH - Bosnia and Herzegovina, BUL - Bulgaria, CRO - Croatia, CYP - Cyprus, CZR - Czech Republic, DEN - Denmark, EST - Estonia, FIN - Finland, FRA - France, GEO - Georgia, GER - Germany, GRE - Greece, HUN - Hungary, ICE - Iceland, IRE - Ireland, ISR - Israel, ITA - Italy, KAZ - Kazakhstan, KYR - Kyrgyzstan, LAT - Latvia, LIT - Lithuania, LUX - Luxembourg, MAL Malta, MON - Monaco, MTN - Montenegro, NET - Netherlands, NOR - Norway, POL - Poland, POR - Portugal, MOL - Republic of Moldova, ROM - Romania, RUS - Russian Federation, SMA - San Marino, SER - Serbia, SLO - Slovakia, SVN - Slovenia, SPA - Spain, SWE - Sweden, SWI - Switzerland, TAJ - Tajikistan, MAC - TFYR Macedonia, TUR - Turkey, TUS - Turkmenistan, UKR Ukraine, UK - United Kingdom, UZB - Uzbekistan.

Table 2. Correlation coefficients in countries of European Region of the World Health Organization, grouped according to GDP per capita

Group according to GDP per capita	Correlation coefficient (r)
GROUP 1: GDP up to 2999 US$	0.182
GROUP 2: GDP 3000-7999 US$	-0.714
GROUP 3: GDP 8000 US$ or higher	-0.190

Interpretation of the results

From results of correlation analysis for three groups according to GDP could be concluded that GDP per capita is an important factor in reducing general mortality in the population, but only in a specified interval. Our results indicate that this interval is approximately between 3000 and 8000 US$. Only when a country attains a certain level of GDP, it could expect that mortality could start to decrease. We could estimate that this threshold is about 3000 US$ per capita. It is interesting that also after a specified threshold increase in GDP per capita, it has no longer influence on reduction of general mortality of a population. These results are valid only for a population level, and are not applicable at an individual level.

Where were in 2001 PH-SEE Network Countries?

The scatter plot presented in Figure 3 could be supplemented with additional information. This could be very illustrative.

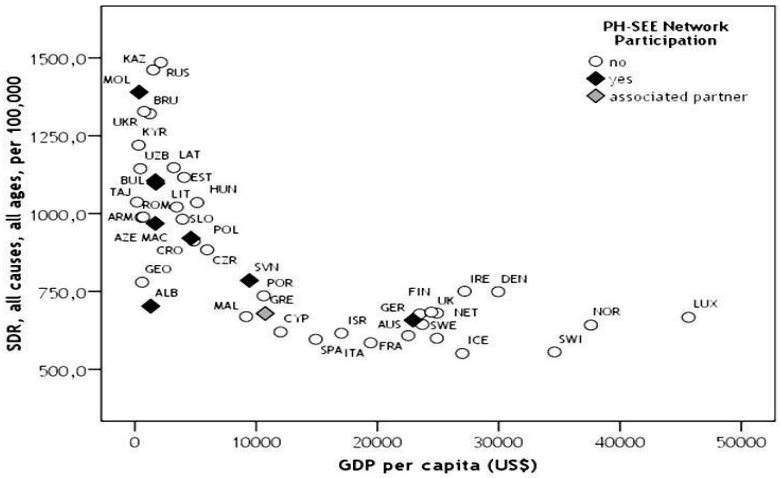
Since we are preparing this module for helping public health teachers from PH-SEE Network, it could be interesting to add on the scatter plot the information on participation of a country in the PH-SEE Network (Figure 4) (18).

Unfortunately, although we have chosen the year 2001 as the year with the most data available on both analyzed variables, data for both variables were not available for Bosnia and Herzegovina, Montenegro, and Serbia for this year in Health for All database for the year 2007.

Some tips for SPSS users

Performing ecological analysis in the SPSS statistical programme is rather simple, nevertheless some tips could be of help.

Figure 4. Standardized death rate (SDR) from all causes in all age groups per 100,000 population in relation to Gross domestic product (GDP) (in US$), in countries of European Region of World Health Organization, by participation in PH-SEE Network (data for 2001)

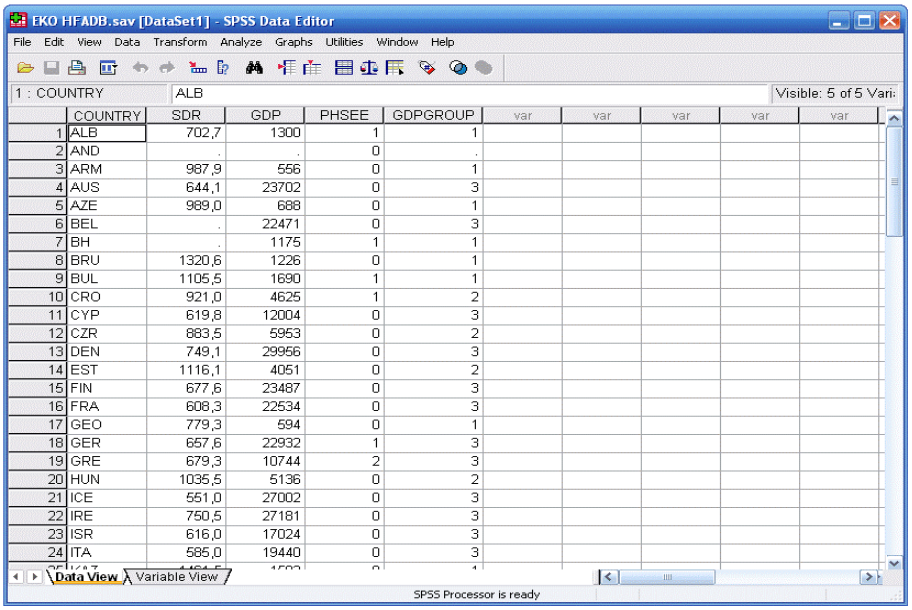


LEGEND: ALB - Albania, BUL - Bulgaria, CRO - Croatia, GER - Germany, GRE - Greece, MAC - Macedonia, MOL - Moldova, ROM - Romania, SVN - Slovenia.

Data entry

Figure 5 presents the data matrix for analysis of ecological study data.

Figure 5. SPSS Data Editor window with data properly prepared for making ecological study scatter plot in SPSS statistical programme



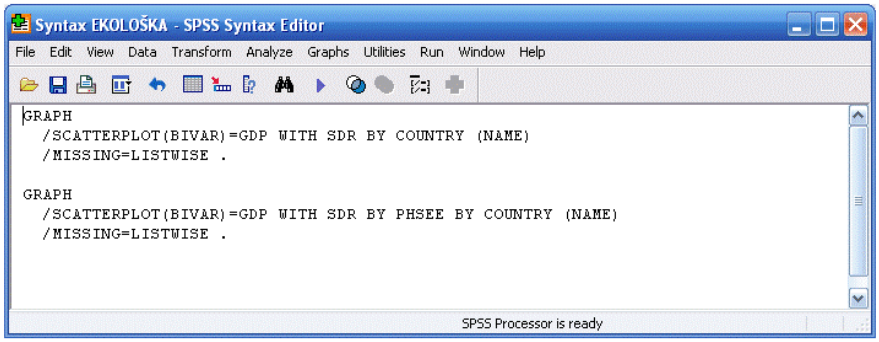
	COUNTRY	SDR	GDP	PHSEE	GDPGROUP
1	ALB	702,7	1300	1	1
2	AND	.	.	0	.
3	ARM	987,9	556	0	1
4	AUS	644,1	23702	0	3
5	AZE	989,0	688	0	1
6	BEL	.	22471	0	3
7	BH	.	1175	1	1
8	BRU	1320,6	1226	0	1
9	BUL	1105,5	1690	1	1
10	CRO	921,0	4625	1	2
11	CYP	619,8	12004	0	3
12	CZR	883,5	5953	0	2
13	DEN	749,1	29956	0	3
14	EST	1116,1	4051	0	2
15	FIN	677,6	23487	0	3
16	FRA	608,3	22534	0	3
17	GEO	779,3	594	0	1
18	GER	657,6	22932	1	3
19	GRE	679,3	10744	2	3
20	HUN	1035,5	5136	0	2
21	ICE	551,0	27002	0	3
22	IRE	750,5	27181	0	3
23	ISR	616,0	17024	0	3
24	ITA	585,0	19440	0	3

We need to have five variables, which means five columns. Beside the columns with data on SDR and GDP, we should have at least the column with country codes. Additionally, we provided the information on participation in PH-SEE Network (codes: 0 - no, 1 - yes, 2 - associated partner), and group according to GDP (codes: 1 - GDP up to 2999 US$, 2 - GDP 3000-7999 US$, 3 - GDP 8000 US$ or higher).

Scatter plot

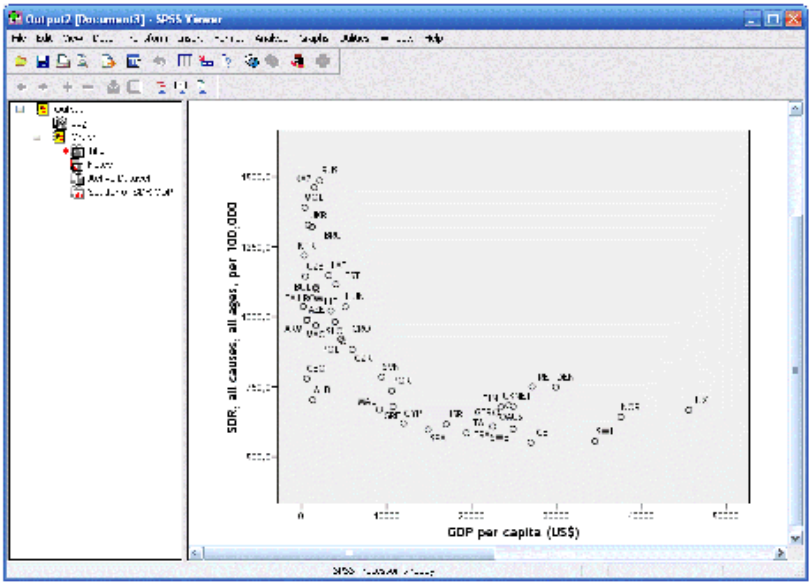
It is rather complicated to explain how to draw a scatter plot by using options offered by menu Graphs that is available for example from the SPSS Data Editor window. Instead of this we are rather providing SPSS Syntax Editor window with syntax for making ecological study scatter plots as presented in Figures 3 and 4. The syntax is provided in Figure 6.

Figure 6. SPSS Syntax Editor window with syntax for making ecological study scatter plots as presented in Figures 3 and 4 in SPSS statistical programme



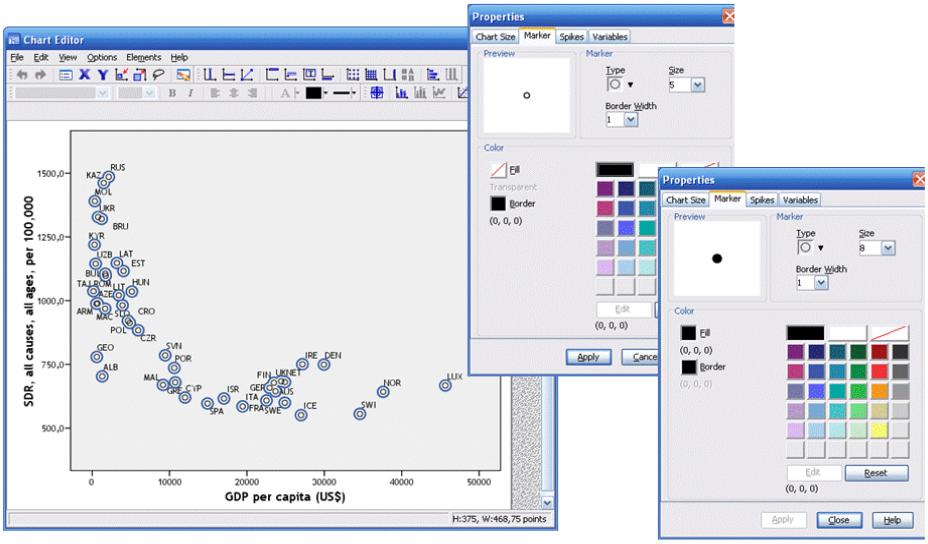
The results of running the first syntax presented in Figure 6, is provided in Figure 7.

Figure 7. SPSS Viewer window with basic scatter plot in SPSS statistical programme



Basic (default) scatter plot could be adapted according needs of the use using SPSS Chart Editor.

Figure 8. SPSS Chart Editor window with communication windows for adapting scatter plot according to user needs (e.g. as presented in Figures 2 and 3) in SPSS statistical programme



Correlation analysis

Correlation analysis could be performed in SPSS using different procedures. The most simple is to use procedure Bivariate Correlation (Figure 9). The results of running this procedure are presented in Figure 10.

Figure 9. SPSS dialog box for running the Bivariate Correlation procedure in SPSS statistical programme

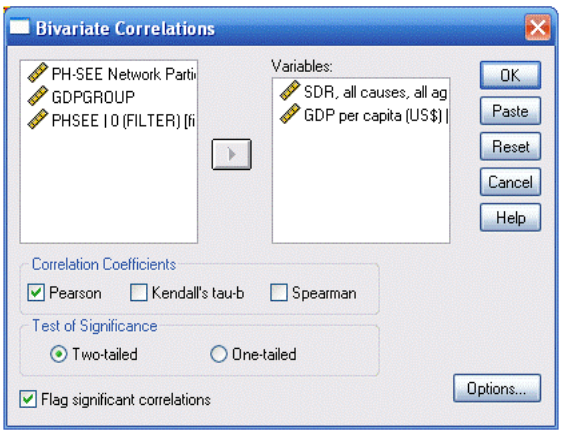


Figure 10. SPSS Viewer window with results of running Bivariate Correlation procedure in SPSS statistical programme

Correlations[a]

		SDR, all causes, all ages, per 100,000	GDP per capita (US$)
SDR, all causes, all ages, per 100,000	Pearson Correlation	1	,182
	Sig. (2-tailed)		,516
	N	15	15
GDP per capita (US$)	Pearson Correlation	,182	1
	Sig. (2-tailed)	,516	
	N	15	18

a. GDPGROUP = 1

Correlations[a]

		SDR, all causes, all ages, per 100,000	GDP per capita (US$)
SDR, all causes, all ages, per 100,000	Pearson Correlation	1	-,714*
	Sig. (2-tailed)		,047
	N	8	8
GDP per capita (US$)	Pearson Correlation	-,714*	1
	Sig. (2-tailed)	,047	
	N	8	8

*. Correlation is significant at the 0.05 level (2-tailed).

a. GDPGROUP = 2

Correlations[a]

		SDR, all causes, all ages, per 100,000	GDP per capita (US$)
SDR, all causes, all ages, per 100,000	Pearson Correlation	1	-,190
	Sig. (2-tailed)		,410
	N	21	21
GDP per capita (US$)	Pearson Correlation	-,190	1
	Sig. (2-tailed)	,410	
	N	21	22

a. GDPGROUP = 3

Case study 2: Hypertension (self-rated) in relation to average monthly gross earnings per person in paid employment in nine health region of Slovenia

Introduction

The second case study is basing on Slovene data. In this case we have used data from two different sources, being:

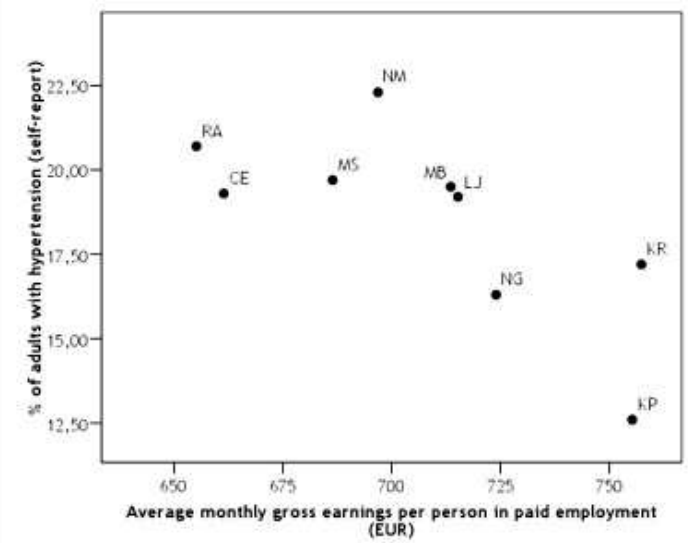
- data from the Statistical Yearbook of the Statistical Office of Republic of Slovenia (19) on gross earnings per person in paid employment on the community level. Since we needed the information on the level of nine health regions, we averaged values of communities, covered by corresponding Regional Institutes of Public Health;
- data on hypertension prevalence from CINDI Health Monitor Survey 2001 (20).

Thus we used data from one routine data source and data from one periodical survey.

Scatter plot

In Figure 11, the scatter plot of hypertension prevalence, and gross earnings per person in paid employment, is presented.

Figure 11. Percent of adults with hypertension (self-reported) in relation to average monthly gross earnings per person in paid employment in EUR, in nine health regions of Slovenia (data for 2001)



LEGEND: CE - Celje Health Region, NG - Nova Gorica Health Region, KP - Koper Health Region, KR - Kranj Health Region, LJ - Ljubljana Health Region, MB - Maribor Health Region, MS - Murska Sobota Health Region, NM - Novo mesto Health Region, RA - Ravne Health Region.

From this chart we could assume that a rather strong negative correlation is present between variables we put in relation. The relationship is very close to linear, thus Pearson's correlation coefficient could be calculated.

Correlation analysis

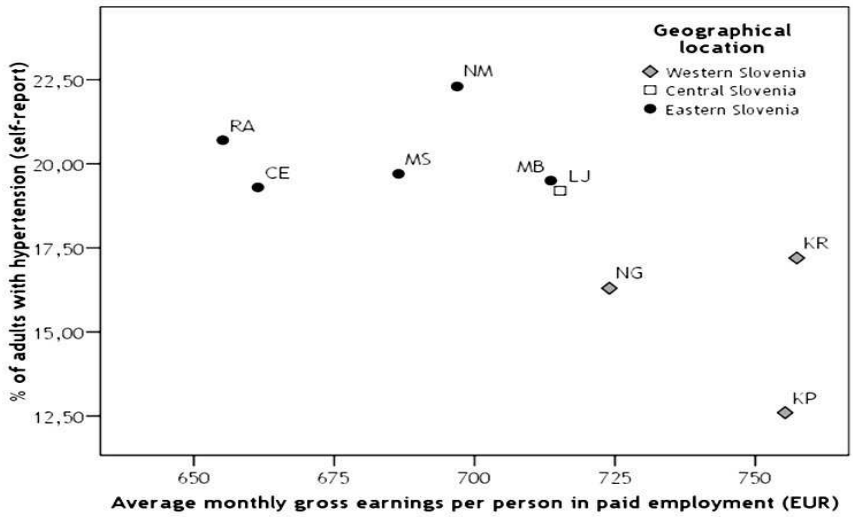
Correlation analysis confirmed our observation – value of Pearson's correlation coefficient is – 0.716, indicating a rather strong negative correlation.

Interpretation of the results

From results of correlation analysis could be concluded that amount of average monthly gross earnings per person in paid employment is an important factor in reducing hypertension in the population.

Again, these results are valid only at a population level, and in no case at an individual level.

Figure 12. Percent of adults with hypertension (self-reported) in relation to average monthly gross earnings per person in paid employment in EUR, in nine health regions of Slovenia by rough geographical location of health regions (data for 2001)



LEGEND: CE - Celje Health Region, NG - Nova Gorica Health Region, KP - Koper Health Region, KR - Kranj Health Region, LJ - Ljubljana Health Region, MB - Maribor Health Region, MS - Murska Sobota Health Region, NM - Novo mesto Health Region, RA - Ravne Health Region.

Is there some other pattern?

The scatter plot presented in Figure 11 could be supplemented with additional information on geographical location of health regions. This could be very illustrative (Figure 12).

It is interesting to note that the situation is the best in the western part of Slovenia, and it is worsened in the direction towards the eastern part of the country.

Exercises

Task 1

Carefully read the theoretical background on this module, and recommended readings.

Task 2

Using the snowball technique, discuss the characteristics of ecological studies. Pay special attention on the problems in interpretation of results of this type of epidemiological studies.

Task 3

In table 3, you will find data on obesity prevalence for 12 statistical regions of Slovenia for the year 2001. From the Web Page of Statistical Office of Republic of Slovenia (http://www.stat.si/eng/index.asp) find Statistical Yearbook with corresponding data (http://www.stat.si/eng/pub_letopis_prva.asp) on gross earnings per person in paid employment for 12 statistical regions of Slovenia (NOTE: in 2001 in Slovenia the currency was Slovenian tolar; the conversion rate to Euros is 1: 239.64). If available, make the scatter plot using SPSS statistical programme, otherwise make it manually.

Table 3. Data on obesity prevalence for 12 statistical regions of Slovenia for the year 2001. Data originate from CINDI Health Monitor Survey 2001 (20), and were prepared exclusively for this module

Statistical region	Obesity prevalence (%)
1. Pomurska	18.8
2. Podravska	16.3
3. Koroska	11.2
4. Savinjska	16.1
5. Zasavska	18.8
6. Spodnjeposavska	21.6
7. Jugovzhodna Slovenija	17.8
8. Osrednjeslovenska	13.2
9. Gorenjska	12.6
10. Notranjsko-kraska	14.4
11. Goriska	9.6
12. Obalno-kraska	14.0

Task 4

In a group of three students prepare an example of ecological study using the World Health Organization, Regional Office for Europe "Health for all Database" (http://www.euro.who.int/hfadb). Make a choice by yourselves. Provide a short interpretation. The results are meant to be part of an assessment.

References

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maindenhead: Open University Press; 2005.
3. Gail MH, Benichou J. Encyclopedia of epidemiologic methods. Chichester: John Wiley&Sons; 2000.
4. Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. Environ Health Perspect 1993; 110:23-38.
5. Armstrong BK, White E, Saracci R. Principles of Exposure Measurement in Epidemiology. Monographs in Epidemiology and Biostatistics, New York Oxford University Press; 1994.
6. McMichael AJ. Setting environmental exposure standards: The role of epidemiologist. Int. J Epidemiol 1989;18:10-16.
7. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. p.231-262. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
8. Hennekens CH, Buring JE. Epidemiology in medicine. Boston, Little, Bown and Company; 1987.

9. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders company; 2001.
10. Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. Fourth edition. London: BMJ Publishing group, 1997.
11. Kirkwood B, Sterne JAC. Essential medical statistics. Second edition. Oxford: Blackwell Publishing Ltd.; 2005.
12. Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1993.
13. Petrie A, Sabin C. Medical statistics at a glance. Malden, Mass: Blackwell; 2005.
14. Morgenstern H. Uses of ecologic analysis in epidemiologic research. Am J Public Health 1982;72:1336-44.
15. Walter SD. The ecologic method in the study of environmental health. Overview of the method. Environ Health Perspect. 1992;94:61-65.
16. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health. 1995;16:61-81.
17. World Health Organization, Regional Office for Europe. European Health for all Database, HFA-DB. Copenhagen: World Health Organization, Regional Office for Europe, 2007. Available from URL: http://www.euro.who.int/hfadb. Accessed: June 4, 2013.
18. Forum for Public Health in South Eastern Europe, Programmes for Training and Research in Public Health. PH-SEE Participating Centres Page [Homepage on the Internet]. Zagreb: Andrija Stampar School of Public Health, 2009 Aug 9. Available from URL: http://www.snz.hr/ph-see/centres.htm. Accessed: June 4, 2013.
19. Statisticni urad Republike Slovenije. Statistical yearbook 2006. Ljubljana: Statisticni urad Republike Slovenije, 2006. Available from URL: http://www.stat.si. Accessed: June 4, 2013.
20. Zaletel-Kragelj L, Fras Z, Maucec Zakotnik J, editors. Risky behaviours related to health and selected health conditions in adult population of Slovenia: results of Slovenia CINDI Health Monitor Survey 2001 (in Slovene). Ljubljana: CINDI Slovenia; 2004.

Recommended readings

1. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization; 1998.
2. Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. Fourth edition. London: BMJ Publishing group; 1997.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 4, 2013.
4. Morgenstern H. Uses of ecologic analysis in epidemiologic research. Am J Public Health 1982;72:1336-44.
5. Schwartz S. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. Am J Public Health. 1994;84:819-24.
6. Susser M. The logic in ecological: I. The logic of analysis. Am J Public Health. 1994;84:825-9.
7. Susser M. The logic in ecological: II. The logic of design. Am J Public Health. 1994;84:830-5.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Cross-sectional studies**
Module: 2.13	ECTS (suggested): 0.25
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Ivan Erzen, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Lijana Zaletel-Kragelj Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Cross-sectional study, epidemiologic study.
Learning objectives	After completing this module students should: • know the definition and characteristics of cross-sectional studies; • be familiar with designing phase of cross-sectional studies; • be familiar with planning phase of cross-sectional studies.
Abstract	Cross-sectional studies are observational epidemiological studies of health status of the population in which a cross-section through frequency and characteristics of health outcomes and other health related events like exposures are studied and, therefore, provide prevalence data. They are very applicable in searching for general insight in health states and conditions that last a relatively long time, as well as various risk factors for different diseases. They provide descriptive information for designing other types of epidemiological studies. The module is describing principles of cross-sectional surveys, especially their designing and planning phase.
Teaching methods	An introductory lecture gives the students first insight in characteristics of cross-sectional studies. The theoretical knowledge is illustrated by a case study. After introductory lectures, students first carefully read the recommended readings. Afterwards, students discuss the characteristics of cross-sectional studies with other peers, especially the designing and planning phase of this type of epidemiological studies. In continuation, they need to find published materials (e.g. papers) on cross-sectional studies and present their findings to other students.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire.

CROSS-SECTIONAL STUDIES

Lijana Zaletel-Kragelj, Ivan Erzen

Theoretical background

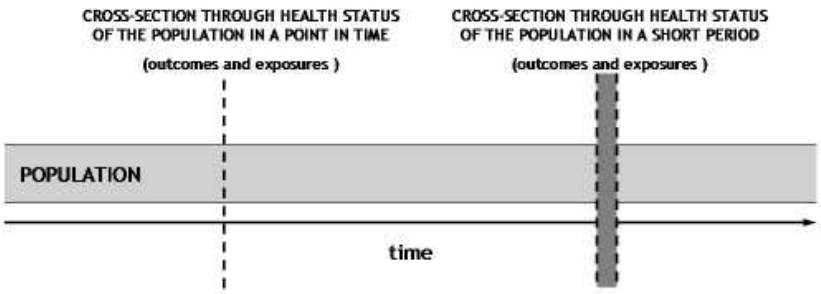
About cross-sectional studies

Definition and description

There exist several similar definitions of cross-sectional studies (CSS):

- according to Last et al. (1), CSS are studies that examine the relationship between diseases or other health-related characteristics, and other phenomena of interest in a defined population at a particular time;
- a summary of several other definitions is that CSS are observational epidemiological studies of health status of the population in which a "snap-shot" of or a cross-section through frequency and characteristics of health outcomes and other health related events like exposures are studied (2-6). This characteristic also gave the name to this type of epidemiological studies;
- CSS are studies that measure the prevalence of health outcomes or determinants of health, or both, in a population at a specific point in time, or over a short period (Figure 1) (4).

Figure 1.Schematic presentation of a cross-sectional study



Health outcomes and other health related events could be measured in CSS on different measurement scale. In those CSS in which the outcome event is dichotomous, the prevalence of this dichotomous event is recorded. This is the reason that CSS are also called prevalence studies (3,5,7,8). Prevalence studies thus could be on one hand regarded as a subgroup of cross-sectional studies, while on the other hand all CSS could be regarded as prevalence studies since we can dichotomize values of every observed outcome.

The selected specific point in time could be a time window within which data are collected (e.g. calendar week or month). It could also be a specific point in time in the course of events, differing in respect of each individual study subject with regard to the actual time (beginning of schooling, retirement, etc.) (3,9,10).

Frequently, CSS are designated as surveys.

Aims

The aims of the CSS are:

- to describe the frequency and characteristics of the observed health-related phenomena at a certain point in time or within a time window (3,6);
- to analyze the relationship between two or more health related phenomena (2,4).

Since CSS are useful in description and analysis of health phenomena, they can be classified among descriptive (3,6), as well as among analytical epidemiological studies (2,4). Today, they represent one of the most important tools of evidence-based public health (11).

Methods

Sources of data

Theoretically, the source for CSS is a population. But, since total population is usually hard to reach, a sample is drawn from the population. Not all members of a population or a sample under observation respond to the invitation and take part in CSS. These relations should be clear.

Advantages and disadvantages

There exist several advantages and disadvantages of CSS (2,4,8,10). They are summarized in Table 1.

Table 1. Some advantages and disadvantages of cross-sectional studies

ADVANTAGES	DISADVANTAGES
9. Relatively easy, quick and inexpensive, 10. Good study design for hypothesis generation, 11. Health outcome is measured at one point in time and the exposure may be measured from the same individuals at the same time and/or historical exposure information may be available, 12. Suitable for studying multiple exposures and/or multiple health outcomes, 13. Suitable for assessment of the prevalence of the events, 14. Suitable for estimating overall and specific health events prevalence, 15. Suitable for observation of frequent states of long duration, 16. Particularly suitable for observation of non-fatal diseases, degenerative diseases with no clear point of onset (e.g. Chronic bronchitis), or for examining effects on physiologic variables (e.g. blood pressure, serum glucose etc.), 17. Suitable for monitoring of the relationship between permanent and invariable exposures (risk factors) and health outcomes, 18. High generalizability, 19. Often good first step for new study issue in public health, 20. Good for public health programmes and health care service planning.	6. Study design not always appropriate, 7. Susceptible to selection bias (possibility of high proportion of long term survivors) i.e. individuals who either recover or die from a disease quickly have less of a chance of being included in the disease group, 8. Susceptible to misclassification (e.g. recall) 9. Not suitable for rare diseases/exposures, or diseases/exposures with short duration 10. Not a useful study design for establishing causal relationships because of problems with temporal sequence of data but this problem could be avoided in repeated cross-sectional studies.

i. Population

All epidemiologic studies are based on a particular population. In this respect we need to distinguish between target and source population:

- target population is the population which is to be subject to inference on the basis of the results of the CSS (1,4);
- source population is the group of participants from whom we have collected data (4). It is also called the study population or base population.

Source and target population could be the same.

ii. Sample

Since CSS is usually not possible to be conducted on the total source population, usually a sample is drawn. There are several methods of sampling. In general, they could be classified into two major groups (9):

- probability sampling – this type of sampling is also called random sampling. Types of random sampling procedures are simple random sampling, systematic sampling, multi-stage sampling, stratified random sampling and cluster sampling (4,9,12);
- non-probability sampling - convenience sampling and purposive sampling are the types of non-probability sampling procedures (4,9,12).

In epidemiology, probability sampling is preferred.

The sample size depends on the characteristics of population under observation, on the purpose of the CSS, methods of data collection and data analysis methods.

iii. Respondents

We usually cannot include all residents invited to take part in the CSS (selected in a sample) as some of them could simply not be found and the others refuse to cooperate. Those willing to take an active part in the CSS are representing only a sub-group of the randomly selected sample. They are called respondents or participants.

A lot of effort should be put to reach as high response rate as possible in order to avoid as much selection bias as possible (4,7,9).

Tools and methods of data collection

There exist several tools and methods for data collection in cross-sectional studies.

i. Tools for data collection

Collection of data in CSS could be carried out:

- by the means of questionnaires which enables to pose the same questions in the same way to each respondent in the CSS. The questionnaire should be as short as possible and each question should be well considered (4,9);
- by the means of health examination including diagnostic and laboratory tests.

According to the tools used to collect data, there exist two main types of CSS (13):

- health interview surveys or HIS surveys in which collection of data is carried out only by the means of questionnaires, and;
- health examination surveys or HES surveys which are usually a combination of questionnaires and health examination including diagnostic and laboratory tests.

2. Methods of data collection

In HIS, questionnaires may be communicated to the randomly selected study subjects in three ways: through mail, through personal interview, or through telephone interview. Each of these methods has their own advantages and disadvantages (9), which are summarized in Table 2.

In HES, the contact between participants and research personnel is personal since the health examination is a component part of the CSS. In this type of surveys, also questionnaires are usually communicated to the randomly selected study subjects through personal interview.

Preparing data for analysis and data analysis

Getting data ready for analysis in CSS starts already at drafting the questionnaire where in respect of individual questions the codes for different answers are already predefined. The encoded data are then entered in the data matrix or data spreadsheet. For data entry the widely used spreadsheet programmes may be applied, however the analysis should be carried out by the means of one of the quality programmes specific for statistical analysis of data.

The basic analyses encompass the assessment of the prevalence[9] of phenomena under observation as a frequency measure in CSS studies.

Whenever we also wish to observe the relationship between a disease and a risk factor, the whole group of observed subjects should be divided with respect to exposure to the effect of the risk factor. By the term "risk factor" different characteristics of respondents are indicated, which are already known to be related to their health status, and which should be prevented, or the extent of effect of which should be decreased (1,15). The assessments of prevalence of observed health outcomes for each observed sub-group should be calculated, which will be subsequently compared.

Strength of associations may be assessed in different ways. We can observe the difference between two prevalence risks or rates, or the prevalence risk or rate ratio, as well as the odds ratio (16).

Presentation and interpretation of the results

The results of CSS may be presented in different ways depending on for whom the presentation is intended:

- to the expert public, results are usually presented in the form of articles published in scientific and other journals, or as presentations at different meetings;
- wider lay public is usually informed through the mass media; nowadays, web pages are also frequently applied media.

Interpretation of results should be carried out with all due attention and impact of possible biases and errors, which were eventually done during the designing, planning or conducting of the CSS, and should be taken into consideration. We should be aware, that the prevalence should be interpreted with a lot of caution, taking into account all potential influences on its value. Even if the errors are not the evident explanation for the observed relationship between the two phenomena, possibility of causality should be assessed very carefully.

Course

The course of a CSS should follow the general principles, which are common to most study designs, including CSS studies.

Phases and periods of the course of cross-sectional studies

The whole course of CSS can be divided into different units.

The smallest unit of the CSS course is called phase of the CSS. Phases of the CSS are usually conducted according to the following scheme (Figure 2) (4,7,17,18):

1. Designing phase
2. Planning phase

[9] NOTE: there exist several different types of prevalence measures: prevalence risk, prevalence rate and prevalence odds (14).

3. Organizing phase (preparation for the implementation)
4. Data collection/analysis phase
5. Results dissemination phase (interpretation and presentation of the results).

As regards the duration, phases can vary significantly. With various circumstances taken into account, they can be conducted successively or in parallel, or may be even intertwined among themselves. Intertwinement is usually more explicitly expressed with the CSS of rapid course. The first and the second, and the third and the fourth phase of the CSS are frequently intertwining.

Periods are wider units of CSS course. A single period may consist of one single phase, or of more phases (Figure 2).

There are numerous textbooks on epidemiological methods containing recommendations with regard to designing and planning and management of each individual phase and/or period (4,17-19).

Although all phases/periods of CSS course are important, planning period is the most important and the most sensitive period. If designing and planning the CSS is in the wrong way, the whole CSS could be set on an inappropriate basis, and the deficiencies of this period are very difficult to be eliminated in the later phases of the CSS. In order to avoid as many faults as possible, the course of the CSS must be planned systematically and with all due care. Therefore, in continuation, special attention is given in this module to recommendations for designing and planning a CSS.

i. Planning period of the CSS
   The first two phases of the CSS are included in the planning period – designing and planning phases.
ii. Preparational period of the CSS
   Preparation period includes more phases than the planning period. In addition to both phases of the planning period also the organizing phase of the CSS is included in this period.
iii. Implementation period of the CSS

Figure 2. Phases and periods of the course of a cross-sectional study (NOTE: to make the presentation easier, in the picture, all phases are represented as being of the same duration, which does not correspond to the facts of the practice; also, overlapping of phases and periods is not taken into consideration)

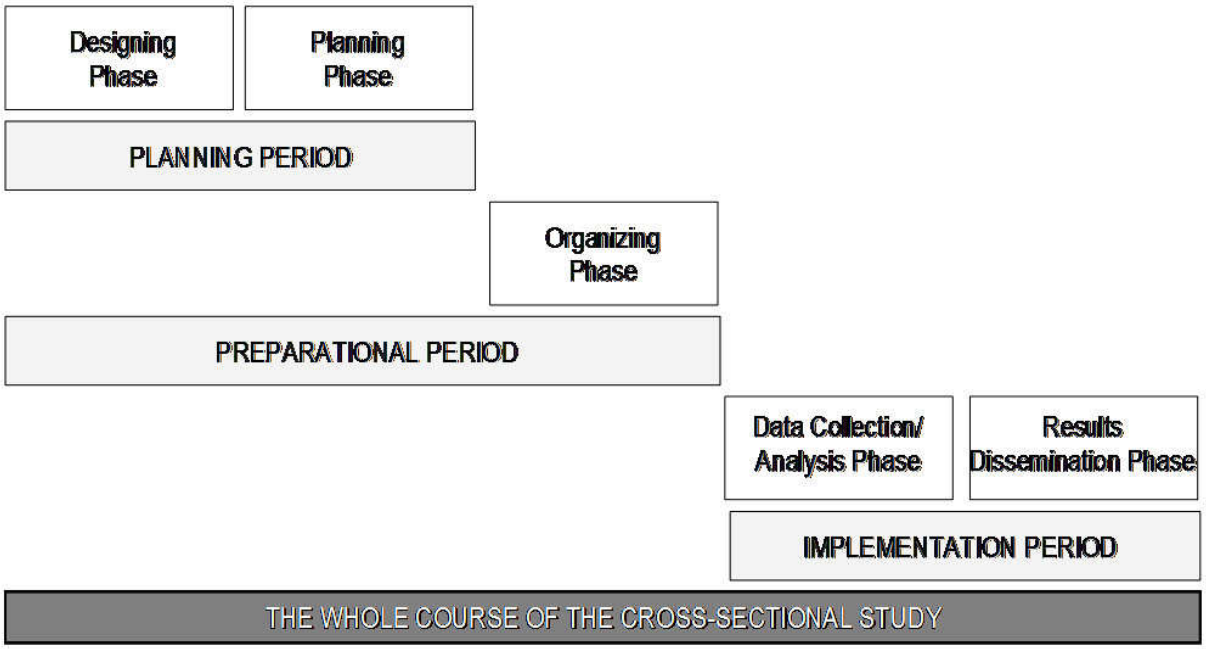


Table 2. Some advantages and disadvantages of three different ways of intermediation of questionnaires to the respondents in cross-sectional studies

ADVANTAGES	DISADVANTAGES
MAIL QUESTIONNAIRES	
1. low costs 2. rapid implementation 3. respondent fills-in the questionnaire at the most appropriate time for her/him 4. high level of anonymity 5. uniformity of posed questions 6. no bias resulting from subjectivity of the interviewer 7. respondent may check her/his answers 8. easy access to respondents	1. possibility of low response rate 2. small individual adaptability at posed questions (no help of interviewer) 3. not possible to use more complicated questions 4. possible influence of social environment (family members) 5. identity of the person who completed the questionnaire can not be controlled 6. order of precedence of answers to questions can not be supervised 7. many questions may not be answered to 8. spontaneity of answers can not be supervised 9. non-verbal messages of the respondent can not be observed 10. possibility of selection bias due to low response rate
FACE-TO-FACE INTERVIEWS	
1. high adaptability of the interviewer to the responder's understanding of questions 2. possibility of high response rate 3. non-verbal messages of respondent can be assessed 4. possibility of supervising the environment in which the respondent is completing the questionnaire 5. possibility of supervising the order of precedence of answering the questions 6. higher possibility of spontaneity of answers 7. possibility to control the identity of respondent 8. possibility of posing more complicated questions 9. supervision over completeness of answers	1. very high costs 2. conduction taking long time 3. respondent can not supervise and check her/his answers 4. possibility of not-suitable time for filling-in 5. low anonymity rate 6. less unified way of posing questions 7. possibility of difficult access to respondents 8. possibility of bias due to the influence of interviewer
TELEPHONE INTERVIEWS	
1. possible adaptability of the interviewer to the responder's understanding of questions 2. possibility of supervising the order of precedence of answering the questions 3. possibility of spontaneity of answers 4. possibility of posing more complicated questions 5. supervision over completeness of answers	1. high costs 2. conduction taking a long time, however not as long as at personal interviews 3. respondent can not supervise and check her/his answers 4. high possibility of not-suitable time for filling-in 5. possibility of low response rate 6. low anonymity rate, however higher than at personal interviews 7. less unified way of posing questions 8. possibility of bias due to the influence of interviewer 9. the control of the identity of person who completed the questionnaire not possible 10. non-verbal messages of the respondent can not be observed

Recommendations for designing a cross-sectional study

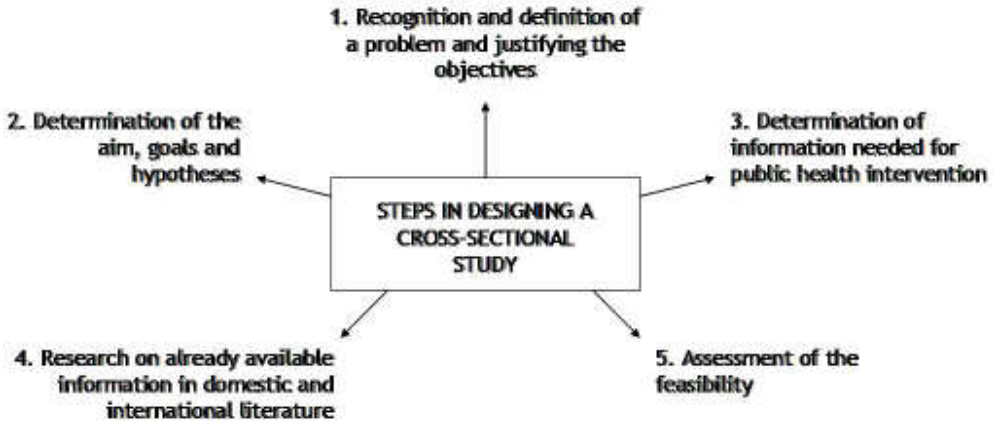
Designing of a CSS is a creative process. A precise management of individual CSS is very difficult to be advised. Nevertheless, common recommendations on actions in this phase of the CSS, do exist (4,10,17,19) and are

very similar to those applied in biomedical studies in general. They can be summarized in the following steps (Figure 3):

i. Recognition and definition of a problem and justifying of the objectives for the CSS

The cornerstone of a CSS is usually a public health event, which is recognized as a problem, which is so noticeable that we would like to investigate it and search for the possible solutions. However, this is not enough. Since the public health research field is so wide, we usually meet with problems already when focusing on the access to a certain problem. Our decision on being interested in something does not suffice for initiation of the CSS. This is only to define the study problem. The selected scope of the CSS must also be grounded. Since CSS are usually associated with big funds, these should be grounded by arguments.

Figure 3. Steps in designing a cross-sectional study



The scope of the CSS is usually grounded on one or more objectives. However, the whole process of activities connected with the beginning of the CSS, is usually initiated by one of them.

ii. Determination of the aim, goals and of the hypotheses of the CSS.

To have a clear idea on the aims and goals it is one of the most important parts of the designing phase of CSS course. It could be helpful to know that:

- the aim of the CSS is defined as "what we shall strive for" in the CSS;
- the goal as "what should be attained" during the CSS in order to realize our striving to the maximal possible extent;
- hypothesis or assumption is our proposal for understanding of the events and processes (our opinion on connectedness between the events under study).

The aim and the goals should be set clearly. In the opposite case, it could be seemed that the CSS being conducted only as its own purpose.

When determining hypothesis at CSS, problems are usual to arise. CSS are distinguishable from other studies in medicine for being first of all oriented in studying the size of health problems (its extension), rendering the hypothesis often possible to be set only upon the results obtained. CSS are primarily meant to set as a result the hypotheses which should be then tested in another study design.

iii. Determination of information needed for public health intervention

With the aim to collect as much information on the problem as possible and to avoid collecting too much data, we should consider carefully which data to be of use for solving the problem. These are data, among others, for exact evaluation of the size and spread of the problem, as well as data for forming the strategies for solving the problem.

iv. Research on already available information in domestic and international literature

Aiming at determining the problem exactly, as well as the purpose and the goals of the CSS thereof, we have to carry out preliminary review on similar CSS at home or abroad, if available. Such a review could serve on one hand in choosing similar methods, but on the other hand it could result even in a decision that the CSS shall not have to be conducted in such a wide extent as it has been anticipated at the beginning. Besides, it could also point out the possibilities and limitations at studying the selected problem.

v. Assessment of the feasibility of the planned CSS

When the planned CSS is established to be worth of conducting, it is to be estimated if:

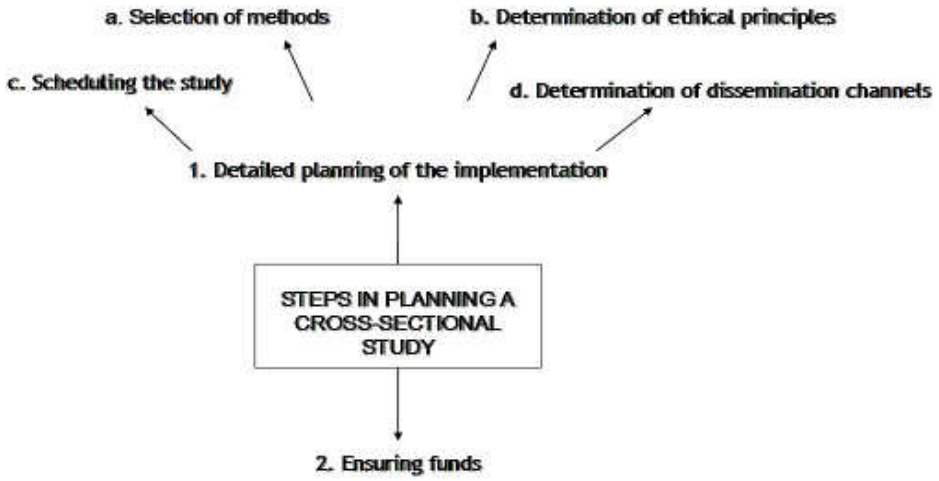
- there are enough funds ensured at least for starting the CSS;
- possibilities are existing for sustainability of the CSS
- there are enough human sources for its execution;
- the CSS is feasible in the population;

- the CSS shall result in the information we are looking for;
- information, when collected and analyzed, will be interesting and useful.

Recommendations for planning a cross-sectional study

When we believe that the CSS is worth of conducting and that at least the main part of it is feasible, we shall puruse with planning and assuring funds for execution (4,17-19). The steps in this phase are as follows (Figure 4).

Figure 4. Steps in planning a cross-sectional study



i. Detailed planning of the implementation of the CSS

Planning of the CSS is considered to be determination of supporting points in conducting of the CSS which shall enable us to make the intended steps, and the way of execution of the study. Planning of the CSS should include the following elements:

a. Selection of the method of approach

Different research methods have been established by different scientific disciplines. The public health research is characterized by intertwining of biomedical and social sciences approaches (6,16). Regardless to this, the following should be determined and/or argued:

- what the target population is
- how to select a sample from the source population and its size
- tools and methods of collection of data
- tools and methods of data analysis
- results dissemination channels.

b. Determination of ethical principles

The question whether to conduct a CSS depends also a great deal on ethical problems. Therefore, ethical instructions for the CSSs have been prepared by the World Health Organization (WHO) (20), basing these on the Helsinki Declaration (21) and other ethical principles (22).

In general, ethical principles of the CSS should stick to basic ethical principles in biomedicine, however differing from them. The CSS of such kind shall bear the responsibility on two levels – on the level of individuals and on the level of the community thereof. Such responsibility does not refer only to individuals and communities taking part in an individual study, but also to other people, whose health could be protected or improved by the means of the results of the CSS. Therefore, all respondents must be informed that with taking part therein something useful shall be done in respect to their own health as well in respect of the health of the community (5,22).

According to ethical principles, each individual should be acquainted with the purpose and course of the study. This should be paid special attention to when possible harmful effects are included. Each individual who decides to cooperate in the CSS shall have the right to withdraw at any time. If the person decides to cooperate we shall ask him/her for his/her consent. This could be in the form of a signed consent - informed consent, or a returned completed questionnaire.

In CSS, the right to privacy of an individual and to confidentiality of information should be respected. However, we are also obliged to inform the population on what we shall be doing and why and which shall be their benefits resulting thereof. Therefore, all proposals for the CSS of this kind should be first addressed to the corresponding commissions (10).

c. Scheduling the CSS

Duration of the CSS should be planned in advance. The timetable of the course on the whole and separate phases must be done as well. In order to be able to follow the timetable, each phase should set dates called milestones.

d. Determination of dissemination channels

In the case of a CSS, also the dissemination channels (i.e. how to present results to different audiences) is recommended to be determined.

ii. Searching for possibilities of ensuring funds for conduction of a CSS

During the planning phase of a CSS, special attention should be given to ensuring the required funds. Since epidemiological CSS are by definition expensive, financing by the means of one big or more sources shall be needed. Therefore, institutions should be found prior to start conducting the CSS, which shall have interest in the problem and which shall also be willing to invest some funds in the implementation.

At the end of this phase, considered also the end phase of the planning period of the CSS, a protocol of the CSS needs to be fully worked out.

The protocol – a plan of a cross-sectional study

Final product of the planning period is a study plan, composed of a conceptual and implementing part. Such plan is called protocol of the CSS (4,7).

The CSS protocol should be composed of the following parts:

- Problem under observation and objectives of the CSS
- Aims and goals of the CSS
- Methods
- Ethical principles
- Time schedule of the course
- Results dissemination channels
- References

Case study

Cindi health monitor survey in Slovenia, 2001

Introduction

The CSS entitled "Risk factors for non-communicable diseases in adults in Slovenia" (23) was the first CSS conducted in Slovenia to such large extent, aiming at studying health behaviour of the population (smoking, nutritional and exercise habits, habits connected to drinking alcoholic beverages, behaviour connected with the road-traffic safety, etc.) on the national and at the same time on the regional level.

By the means of this CSS, an attempt to establish the system of monitoring of health behaviour of the Slovene population was intended to be conducted.

Conceptually, this CSS represented a part of a wider WHO project CINDI Health Monitor (CHM) (24). The project was conducted within the international programme for combating the non-communicable diseases of the WHO Countrywide Integrated Non-communicable Diseases Intervention (CINDI) Programme (24-28).

The ongoing WHO CHM project is mostly aiming at monitoring, assessing and comparing the trend of health behaviour in CINDI countries with different politically-economic systems. Owing to comparability, monitoring should be conducted under a uniform methodology.

The CHM methodology is based on the project Finbalt Health Monitor (24,29), which involves Finland and Baltic Republics Lithuania, Latvia and Estonia. The project is coordinated by Finland, due to its rich tradition and great experiences and successes in preventing non-communicable diseases. The most well-known project is the North Karelia Project (24). In this project, which is now lasting for more than 25 years, Finns succeeded to decline the coronary mortality rate among the male population aged 35-64 years, by about 75% (30). Besides, Finland is collecting data on health behaviour of its adult population on a random sample each year, starting in 1974. For this purpose they make use of mail CSS. Through the project Finbalt Health Monitor they started to spread their experiences also to the neighbouring countries. In 1990, Estonia was first to joint the project, to be then followed by Latvia in 1994 and Lithuania in 1998.

Since the Slovene survey of health behaviour represents a part of the CHM project, it can be also called CHM survey of Slovenia (CHMS-SI).

The activities associated with the CHMS-SI started at the end of February 2001. At that time, a first joint meeting was organized upon the initiative of the representatives of the Ministry of Health (MH). The following representatives and institutions participated, the knowledge and capacities of which enabled them to collaborate in the survey: CINDI Slovenia (CINDI-SI), Faculty of Medicine (FM), National Institute of Public Health of the Republic of Slovenia (NIPH) and regional health protection institutes (RHPI).

The first two tasks of invited public health experts were to communicate the idea of the survey to their institutions, and to assess the feasibility for carrying out such an extensive survey in Slovenia, which was to result in setting the basis for the system of monitoring of the health behaviour of the population. The estimation was not easy

to be reached. The biggest problem related to an extremely short term planned for the survey to be started. It was planned to be started in the middle of May of the same year. The third task was to find other public health experts, capable and willing to cooperate in such an extensive project outside the invited group. The meeting was aiming at forming a research group in the shortest possible time, which would be able to make the entire plan for the survey within the available period of time, and also to initiate the implementation thereof within the planned term.

All institutions which were invited to collaborate in the project, agreed to collaborate. The research group was also formed quickly. Within the short time given, the group managed to make the plan for implementation of the survey and to start the survey within the planned term.

Designing phase

Recognition and definition of a problem and justification of the objectives of the survey

The problem

In Slovenia, the problem of preterm mortality caused by non-communicable diseases has become so important that the need for studying and solving it has become evident (23,31).

Non-communicable diseases in Slovenia were, similar to the other European countries, one of the leading causes of death (31). Mortality rate resulting from such diseases was in Slovenia higher than for example in France, Germany, Italy, or in Finland, but lower than in the Baltic republics (31).

Objectives

The need for studying and solving the problem of preterm mortality, caused by non-communicable diseases, was based on following objectives:

i. The first objective was related to the risk factors for non-communicable diseases.
The problem of non-communicable diseases is mostly related to behavioural risk factors (12). When reviewing the existence of studies conducted within the period from 1991 on, the real estimation of the extension of these factors in Slovenia was shown not to be actually assessed before the CHMS-SI was started (32).

ii. Second objective represented the guidelines of the Slovene health policy in the field of health care and health insurance development. These guidelines were determined in the National Health Care Programme of the Republic of Slovenia (NHCPRS) (33).
In this programme, rising of the quality of health of the Slovene inhabitants was pointed out, as well as adjustment and improvement of the health system's operating in relation to the financial possibilities of Slovenia. In the programme, also, the aims of the WHO were taken into consideration, which have been laid down in two documents: "Health for all" (34) and "Health 21" (35). The priority tasks of the NHCPRS were among others the following:
    - stimulating all stakeholders of the health care to collaborate in forming and carrying out of the programmes on health promotion;
    - stimulating research of the population health by the means of interdisciplinary researches, first of all researches of life style and health behaviour patterns of people and the influence thereof on the health of the population;
    - taking part in the comprehensive programmes of health promotion of the WHO and EU, and;
    - implementing the strategy of health protection of the population in compliance with the guidelines laid down in the documents of the WHO.

    All the stated above was jointed in the CHMS-SI.

iii. The third objective was the key aim of the international CINDI programme (23,25-28), based on the health promotion and prevention of the risk factors for non-communicable diseases. Slovenia has been taking part in the CINDI programme unofficially for over ten years, however officially from 1993. Also, the aims thereof were in compliance with the strategic aims of the WHO.

iv. The fourth objective was the new development policy of the CINDI programme.
A decision was adopted in June 2000 at the headquarters of CINDI at the WHO Regional Office for Europe, to upgrade their activity (25-28). Surveys conducted up to that time, were limited only to demonstration areas of the participating countries. There were two surveys already conducted in Slovenia up to that time in 1991 and in 1996 (36,37), which were limited to the Ljubljana region only. With the stated date, the activities were supposed to be moved from community-regional level to the national level. The surveys were to be conducted each two years at most, instead of every five years. The activities were entitled CINDI Health Monitor (CHM) (24). This decision was based on the assessment that behavioural patterns linked with non-communicable diseases in the wider Europe area, needed to be evaluated and improved. At the CINDI Winter School, held about one half of the year later in Helsinki, on 12 and 13 February 2001 (24), the representatives of Slovenia, among whom the representatives of the MH, the NIPH, RHPIs and CINDI-SI, estimated that a survey under the methodology CHM could be conducted in Slovenia already in the spring of 2001. Upon this decision, our survey was started.

Determination of the aim and goals

Aiming at contributing to detailed knowledge on health behaviour of adult population in Slovenia and consecutively contributing to realisation of the measures of the priorities determined in the NHCPRS (33), the following main goals were set (23):

- to investigate smoking habits on national and on regional level;
- to investigate nutritional habits on national and on regional level;
- to investigate alcohol consumption habits on national and on regional level;
- to investigate physical activity habits on national and on regional level;
- to investigate oral health habits on national and on regional level;
- to investigate road safety habits on national and on regional level;
- to investigate the burden of stress on national and on regional level.

The hypotheses were not determined at this phase of the survey since they were to numerous. It was decided that detailed hypotheses should be the issue of specific studies based on CHMS-SI database.

Definition of what information is needed

In the case of a CHMS-SI survey, we did not have to consider which information should be the result of the survey in order to be able to assess the size of the problem and to find the ways for the resolution thereof, since with entering the project CHM we accepted a common international questionnaire (24). The task of research group was to translate the questionnaire correctly into the Slovene language and to adjust it to conditions in Slovenia.

Research on potential availability of the required information

The review of studies dealing with the prevalence of risk factors for non-communicable diseases in Slovenia (32) was carried out prior than the initiative for the survey has been started. In the review, the results of studies were included, which were carried out in Slovenia after 1991 when our country became independent. These were the results of surveys of the programme CINDI-SI, of the Institute of Oncology, a survey within the framework of the Slovenian public opinion (SPO), and some other surveys (36-45). By the means of this review, the urgent need for exact evaluation of the spread of risk factors for non-communicable diseases in Slovenia as a whole and in each individual health region was established.

Assessment of feasibility of the planned survey

When assessing the feasibility of the planned survey, the funds showed to suffice at least for the beginning. Human resources needed for the start of conducting of the survey were not questionable as well. Moreover, the information, if possible to be collected and analyzed, was seemingly applicable and interesting.

When estimating the feasibility of the survey, the response of the Slovene population to the survey and feasibility due to a short time available for the preparations thereto, turned out to be two major unknown items.

Regarding the response of the population, studies we were able to find in the short limited term in relation to the estimation of the response (46-48) indicated that, in Slovenia all kinds of response rates could be expected, from very good to very bad. Therefore, we were also not able to estimate whether we would acquire the searched information.

Nevertheless, the preparations were carried on, since the CHMS-SI conducted in 2001 should have represented also the study of feasibility of setting the monitoring system of the health behaviour of the adult population in Slovenia.

Planning phase

Planning of the CHMS-SI survey

The planning of the CHMS-SI was in great extent directed by the WHO CHM international project. Recommendations within the framework of this project were the following (24):

i. Organisational recommendations

Organisational recommendations are as follows:

- the survey should be conducted for the first time during the period 2001-2002, if possible,
- the survey should be conducted in all countries at approximate the same season, recommended period being March-May,
- it should be carried out, if possible, at the national level,
- it should be carried out, if possible, every second year.

ii. Methodological recommendations

Methodological recommendations are as follows:

- the sample should be selected upon the principles of simple random sampling,
- it should be of the size of at least 3000 units,
- target population should be adults preferably of the age between 25-64 years (the range could also be wider with regard to the needs of individual country),
- the survey should be based on a common core WHO CHM questionnaire,
- a self-administered postal questionnaire is recommended; if not possible, face-to-face or telephone

interviews can be used,

- non-respondents should not be replaced with other individuals. They should be reminded by sending them a new invitation.

Exact planning of the course of the survey

Selection of methods

When planning survey methods in Slovenia, we mostly aimed at holding on the recommendations of the CHM, but this was not always possible. Whenever another methodology had to be applied, we tried to deviate from the recommended to minimal possible extent.

i. Selection of target population

We defined the target population in compliance with the recommendations of the CHM (24), meaning to have included in the study adult population aged 25-64 years.

If comparing Slovenia with other countries in which similar studies were conducted (30), determination of the target population was characterized by regional approach. Since differences existed in mortality and certain diseases prevalence and incidence, and in some other socio-economic indicators between individual health regions in Slovenia (49-52), we decided to observe the target population separately in respect of each of 9 health regions.

ii. Sampling method and sample size

According to the recommendations (24), the simple random sampling should be applied and the sample, if possible, should be based on the population-based registry. The sample for studying the risk factors was selected on the basis of database of the Central Registry of the Population (CRP) (53), what entirely corresponds to the recommendations. But we selected a slightly different sampling method. Since a single health region was selected as basic observational unit, the stratified random sampling was applied. Individual stratum of the sample was represented by an individual health region of Slovenia. Such sampling method was allowed by the CHM.

When determining the sample size we were subject to the limitation of the smallest number of 3000 units, determined by the CHM. Due to the regional approach, the sample size in Slovenia was differing from that in countries in which a similar study was conducted (30). As per the final estimations, the anticipated number of inhabitants included in the survey was 15,426 for the whole Slovenia (from 578–4591 units in individual regions). We anticipated such high number with the planned multivariate methods of data analysis on regional level and with the planned postal administration of questionnaires due to which also the drop-out (non-response) had to be taken into consideration when estimating the sample size (23).

3. Questionnaire

The questionnaire used in the CHMS-SI was entitled "The Health Behaviour Questionnaire". It originated from the project Finbalt Health Monitor (24,30). The original CHM questionnaire was slightly adjusted to the circumstances in Slovenia.

The content of "The Health Behaviour Questionnaire" was arranged into the following data groups:

- basic demographic data on respondent
- habits of the use of some medical services and evaluation of the respondent's health status
- smoking habits
- nutritional habits
- alcohol beverages consumption habits
- physical activity habits
- road safety habits.

Altogether, there were 73 questions.

4. Method and course of interviewing

Upon proper consideration, we selected the mail survey as the method of interviewing. Such a method has also been recommended by the project CHM (24).

To individuals who would not respond within 14 days, we decided to deliver a reminder accompanied once more by the questionnaire. If the first reminder would still not be responded to, we decided to deliver them another one within one week, consisting of only a letter asking individuals to answer the questionnaire.

If the results of simultaneous analysing of the returning of questionnaires in Slovenia and by its regions would indicate the response to be worse as planned at the beginning, the first reminder would be followed by the permitted methods for stimulation of cooperation. For such purpose, the rewarding with "healthy items" was planned, such as visiting Slovene health resorts, healthy food such as fruit, etc.

Our expectations were that the final response of the respondents in individual regions would be at least 45%.

5. Statistical methods and tools

Besides basic analysis, among which survey of distribution of behavioural risk factors in Slovenia as a whole and by individual health regions, and basic analysis of association, also multivariate methods were planned (23).

In respect of each of the planned analyses, methodological instructions should be prepared and, if necessary, education on use of the instructions in praxis, since analysis should not be carried out only in one institution, but widespread through all health regions.

In respect of all kinds of analyses, the actual version of the statistical programme SPSS (Statistical Package for Social Sciences) for Windows was planned to be used.

Database for analysis was planned to be prepared upon the recommendations of the WHO CHM group (24,54): the same names of variables and the same codes thereof should be applied as proposed by the WHO CHM group.

Determination of ethical principles

In the survey CHMS-SI, ethical principles and provisions of the Personal Data Protection Act (55) were implemented in the following manners:

i. Informing population about the survey (why, when and how it would be conducted and its benefits).
   The communication channels were:
   - mass media - planned to be started about fourteen days prior to sending the questionnaire, and through advertising materials in the form of posters, aiming at getting familiar with the survey of as wide public as possible and not of only the respondents;
   - a cover letter to the questionnaire and the questionnaire itself.

ii. Informed consent for cooperation in the survey.
   An informed consent, a completed questionnaire returned by the respondents, was also considered.

iii. Respecting the right to privacy.
   We intended to respect this right as much as possible, therefore this was also one of the reasons for having decided for a mail survey approach.

iv. Protection of personal data.
   Due to the expected low response rate, we had to anticipate at least one reminder, on account of which we had to keep a record on the respondents already having returned the questionnaires. For this reason, the interviewing could not be conducted absolutely anonymously. Nevertheless, the anonymity was ensured to the maximal possible extent. At carrying out the analysis, names and surnames of individual respondents were replaced by special identification numbers, and only specific persons were authorised to know the connection between IDs and personal data.

v. Application for approval by the authorized ethical commission.
   In Slovenia, a rule is in force, according to which each study from the field of medicine, including epidemiological studies, which does otherwise not interfere with physical integrity of people, should be acquired a consent of the Ethical Commitee of the Republic of Slovenia prior to its conduction. The survey was therefore presented to this ethical body. It was approved at the beginning of April 2001.

Scheduling the survey

The anticipated duration of the survey from the start to the dissemination of majority of the results was three years. During this period, the activities and analyses should have taken place first, which would contribute to improving the course of the next survey of the same type. The analysis mentioned was, for example, the analysis of adequacy and success of the preparations for the survey and the analysis of efficiency of carrying out the interviewing.

When carrying out statistical analyses, we have first of all anticipated data description, to be followed by the univariate and multivariate analyses. Within each phase of the statistical analysis, drafting of the recruitment of procedures was first planned.

Determination of database management and results dissemination channels

Data were planned to be kept at the Institute of Social Medicine of the Faculty of Medicine, Ljubljana, which would be responsible to prepare data in such form so as to be accessible for the wider public.

The survey on behavioural risk factors should be characterised significantly by accessibility of data for public, and by the results thereof to be communicated simultaneously as quickly as possible.

i. Accessibility of data.
   Data on the CHMS-SI are public, meaning that each Slovene citizen may have access thereto providing to respect the rules and under specified conditions.
   The most important rule was that no data referring to the identity of respondents in the survey should be accessible for wider public.
   In order to avoid disputes arisen between the potential users of data, a special group was anticipated to be established to keep the list of proposers and their proposals and to mutually adjust these.

ii. Informing public on the results of the study and their applicability.
   Results of analysis, which would have been carried out by the members of the study group, were intended to be communicated to the inhabitants of Slovenia regularly and simultaneously. For this purpose, the media should be applied, and the more detailed information should be available on the website of the survey.
   Our intention was to inform the expert public on the results of the survey through a series of publications.

With the aim for the database to be managed qualitatively, a body was anticipated to be established (project's council), to exercise control over operation of the database and realisation of its purpose.

Provision of funds for conduction of the survey

In addition to the funds which the CINDI-SI invested in the initiation of the survey, The survey was also financed by the Ministry of Health and the Ministry of Education and Sports of the Republic of Slovenia.

Exercises

Task 1

Carefully read the part on theoretical background of this module. Critically discuss the characteristics of cross-sectional surveys with your colleagues.

Task 2

From domestic (e.g. Biomedicina Slovenica, and COBISS-Cooperative Online Bibliographic System of Slovenia in Slovenia), and/or international bibliographic data-bases (e.g. Medline, PubMed) find out if any other cross-sectional survey has been already performed in your country.

Task 3

If yes, then try to find out its characteristics. If not, try to find an example from other countries (e.g. FINBALT Health Monitor Surveys).

Task 4

Discuss the characteristics, strengths and limitations of the selected survey with your colleagues.

References

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Abramson JH. Cross-sectional studies. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002. p.509-528.
3. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Company; 1987.
4. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
5. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization; 1998.
6. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
7. Premik M. Introduction to epidemiology (in Slovene). Ljubljana: Faculty of Medicine; 1998.
8. Pearce N. A short introduction to epidemiology. Second Edition. Occasional Report Series No 2. Wellington, NZ: Massey University Wellington Campus, Centre for Public Health Research; 2003.
9. Bailey KD. Methods of social research. New York: The Free Press (A Division of Macmillan, Inc.); 1987.
10. Vaughan JP, Morrow RH. Manual of epidemiology for district health management. Geneva: World Health Organization; 1989.
11. Brownson RC, Baker EA, Leet T, Gillespie KN. Evidence based public health. Oxford, New York: Oxford University Press; 2003.
12. Hansen MH, Hurwitz WN, Madow WG. Sample survey methods and theory. Volume I Methods and applications. New York: John Wiley Sons; 1993.
13. Aromaa A, Koponen P, Tafforeau J (uredniki).Status and future of health surveys in the European Union. Final report of Phase 2 of the project Health Surveys in the EU: HIS and HIS/HES Evaluations and Models. Helsinki, Bruxelles: National Health Institute KTL, Scientific Institute of Public Health; 2003.
14. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
15. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
16. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1993.
17. Lutz W, Chalmers J, Lockerbie L, Hepburn W. Health and community surveys. Volume 1. A practical manual for health and community workers. Geneva: IEA, World Health Organization, MacMillan; 1992.
18. Lutz W, Chalmers J, Lockerbie L, Hepburn W. Health and community surveys. Volume 2. A practical manual for health and community workers. Geneva: IEA, World Health Organization, MacMillan; 1992.
19. Aday LA. Designing and conducting health surveys: a comprehensive guide. San Francisco: Jossey-Bass; 1996.

20. Bankowski Z, Bryant JH, Last JM, editors. Ethics and epidemiology: international guidelines. Proceedings of the XXVth CIOMS Conference. Geneva: Council for International Organizations of Medical Sciences; 1991.
21. World Medical Association. World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. Available from URL:

    www.who.int/bulletin/archives/79(4)373.pdf. Accessed: June 4, 2013.
22. Public Health Leadership Society. Principles of the ethical practice of public health. New Orleans: Public Health Leadership Society; 2002. Available from URL:

    http://209.9.235.208/CMSuploads/PHLSethicsbrochure-40103.pdf. Accessed: June 4 2013.
23. Zaletel-Kragelj L. Risk factors for noncommunicable diseases in adults in Slovenia Ljubljana. Survey protocol (in Slovene). Ljubljana: Institute of Public Health, Faculty of Medicine and CINDI-Slovenia; 2001.
24. Prattala R, Helasoja V, Laaksonen M, Laatikainen T, Nikander P, Puska P. Cindi health monitor. Proposal for practical guidelines. Helsinki: National Public Health Institute; 2001.
25. World Health Organization, Regional Office for Europe. CINDI Countrywide Integrated Noncommunicable Diseases intervention Programme. Protocol and guidelines for monitoring and evaluation procedures. Berlin: Springer-Verlag; 1987.
26. World Health Organization, Regional Office for Europe. CINDI 2000. Positioning CINDI to meet the challenges. Copenhagen, World Health Organization, Regional Office for Europe; 1992.
27. World Health Organization, Regional Office for Europe. CINDI 2000. Prevention in primary care. Recommendations for promoting good practice. Copenhagen: World Health Organization, Regional Office for Europe; 1994.
28. World Health Organization, Regional Office for Europe. CINDI 2000. Protocol and guidelines. Copenhagen: World Health Organization, Regional Office for Europe; 1996.
29. Finbalt Health Monitor. Feasibility of a Collaborative System for Monitoring Health Behaviour in Finland and the Baltic Countries. Helsinki: National Public Health Institute; 1999.
30. Puska P, Tuomilehto J. Nissinen A, Vartainen E. The North Karelia project. 20 year results and experiences. Helsinki: National Public Health Institute; 1995.
31. World Health Organization, Regional Office for Europe. The Health for all Statistical Database. Copenhagen: World Health Organization, Regional Office for Europe; 2007.
32. Zaletel-Kragelj L, Pokorn D. Risk factors for diseases in Slovenia (in Slovene). Ljubljana: University of Ljubljana, Faculty of Medicine; 2000.
33. National health care plan of Republic of Slovenia - Health for all to 2004 (in Slovene). Ur List RS. 2000;10:6650-77.
34. World Health Organization, Regional Office for Europe. Targets for health for all. Targets in support of the European strategy for health for all. Copenhagen: World Health Organization, Regional Office for Europe; 1989.
35. World Health Organization, Regional Office for Europe. Health 21. The health for all policy framework for the WHO European Region. Copenhagen: World Health Organization, Regional Office for Europe; 1999.
36. Caks T. Relation of smoking to other risk factors for noncommunicable diseases in Ljubljana population (in Slovene). Master degree thesis. Zagreb: University of Zagreb, School of medicine; 2001.
37. Gradisek A, Soln D, Trsan V, Zakotnik-Maucec J, Preseren N, Kovac M, Caks T, Bulc M, Gabrovsek S, Milohnoja M. Study of risk factors for noncommunicable diseases in Ljubljana (in Slovene). Zdrav Var. 1992;31:71-7.
38. Tos N et al. Slovenia in transition II. Slovene public opinion 1993–1999 (in Slovene). Ljubljana: University of Ljubljana, Faculty for Social Sciences; 2000.
39. Koch V. Some characteristics of nutrition habits of adults in Slovenia (in Slovene). PhD Thesis. Ljubljana: University of Ljubljana, Faculty of Biotechnology; 1997.
40. Moravec Berger D, Ravnikar B, Jezersek P, Lovse B, Pavlin-Klemenc M, Florjancic M, Mikulic C, Gale I. Spread of some known risk factors for cardiovascular diseases in Zgornja Scavnica and Branik (in Slovene). Zdrav Var. 1992;31:63-70.
41. Jezersek P, Accetto R, Cibic B, Dobovisek J, Kolsek B, Lapanja Z, Mihelic-Brcic M, Petrin J, Pirc-Cercek O, Zaletel-Kragelj L, Zemva A. Epidemiological assessment of arterial hypertension in Republic of Slovenia (in Slovene). Zdrav Vestn. 1988;57:403-5.
42. Premik M. Some results of project "Oral health and quality of life in Slovenia (in Slovene). Zdrav Var. 1994,33:85-9.

43. Primic-Zakelj M. Health enlightment of adult Slovenes about cancer (in Slovene). Specialistic thesis. Ljubljana: Institute of Oncology; 1991.
44. Pompe-Kirn V, Japelj B, Primic-Zakelj M. Trends in lung cancer and smoking habits in the male population of Slovenia (in Slovene). Zdrav Vestn. 2000;69:97-9.
45. Pompe-Kirn V, Japelj B. Trends in lung cancer and smoking habits in the female population of Slovenia (in Slovene). Zdrav Vestn. 2000;69:679-82.
46. Prevolnik Rupel V, Rebolj M. The Slovenian VAS tariff based on valutations of EQ-5D health states from the general population. Ljubljana: National Public Health Institute, 2000.
47. Avgustin C. An attempt of measuring of health among adults inhabitants of Celje (in Slovene). 2nd Congress of Slovene Preventive medicine. Book of abstracts. Ljubljana: National health Institute; 1998.
48. Kersnik J. Satisfaction of patients with family doctor (in Slovene). PhD thesis. University of Ljubljana, Faculty of Medicine; 2001.
49. Pompe-Kirm V, Primic Zakelj M, Ferligoj A, Skrk A. Maps of cancer incidence in Slovenia 1978–1987 (in Slovene). Ljubljana: Institute of Oncology; 1992.
50. Cancer incidence in Slovenia 1997 (in Slovene). Ljubljana: Institute of Oncology, Cancer Registry; 2000.
51. Cancer incidence in Slovenia 1998 (in Slovene). Ljubljana: Institute of Oncology, Cancer Registry; 2001.
52. Hanzek M, editor. Report on human development. Slovenia 1998. Ljubljana: Institute for Macroeconomic Analysis and Development; 1998.
53. Trsinar I. Central registry of population (in Slovene). Ur List RS. 1999;9:194.
54. National Public Health Institute, Finland. Finbalt Health Monitor. The data of the 1998 Surveys. Helsinki: National Public Health Institute; 1999.
55. Personal data protection act (in Slovene). Ur List RS. 1999;9:7534-39.

Recommended readings

1. Abramson JH. Cross-sectional studies. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press, 2002. p.509-528.
2. Cross-sectional surveys. In: Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. p.213-229. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 4, 2013.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Health surveys as a powerful tool in planning public health interventions**
Module: 2.14	ECTS (suggested): 0.1
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana Ivan Erzen, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana
Address for correspondence	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	CINDI programme, cross-sectional study, public health intervention programmes, survey.
Learning objectives	After completing this module students should: • know the role of cross-sectional surveys as an effective tool in planning public health interventions; • know the position and importance of cross-sectional surveys in evidence-based public health; • be familiar with some cases of domestic and foreign cross-sectional surveys.
Abstract	Health surveys are observational epidemiological studies of health status of the population in which usually a cross-section through frequency and characteristics of health outcomes and other health-related events like exposures are studied and therefore provide prevalence data. Surveys are very applicable in searching for general insight into the health states and conditions that last a relatively long time, as well as various risk factors for them. Their results could be efficiently used in planning public health interventions, and in fact today they represent one of the most important tools of evidence-based public health The module is presenting basic theoretical background necessary for understanding the usefulness of health surveys in planning public health interventions, as well as it provides a case study.
Teaching methods	An introductory lecture gives the students the first insight into the characteristics of cross-sectional studies. The theoretical knowledge is illustrated by a case study. After introductory lectures, students first carefully read the recommended readings. Afterwards, they discuss the characteristics of health surveys and their potential power for effective health care planning, especially in the field of public health. In continuation, they need to find published materials (e.g. papers) on health surveys and present how they were used (or supposed to be used) in planning public health interventions.
Specific recommendations for teachers	• work under teacher supervision/individual work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire.

HEALTH SURVEYS AS A POWERFUL TOOL IN PLANNING PUBLIC HEALTH INTERVENTIONS

Lijana Zaletel-Kragelj, Ivan Erzen

Theoretical background

Basic definitions and explanations of terms

Surveys and health surveys

Surveys could be defined in several ways, three of them being the following:

- according to A Dictionary of Epidemiology (1), surveys are defined as investigations in which information is systematically collected, but in which the experimental method is not used;
- according to the Encyclopedia of Public Health, surveys are studies used to collect quantitative information about items in a population (2), and;
- according to Rossi and Freeman, surveys are systematic collection of information from a defined population, usually by means of interviews or questionnaires administered to a sample of units in the population (3).

Health surveys are surveys designated to provide information on the health status of a population. They could be descriptive, exploratory, or explanatory. Synonym for a disease frequency survey is a cross-sectional study (1).

In health surveys, data could be collected by the means of questionnaires (face-to-face interview, telephone interview, or self-completed questionnaires), or by the means of health examination, usually in combination with interview. According to which tool is used to collect data in health surveys, there exist two main types of them (4):

- health interview surveys or HIS - surveys in which collection of data is carried out only by the means of questionnaires. In HIS, questionnaires may be communicated to the study subjects in three ways: through mail questionnaire, through personal or face-to-face interview or through telephone interview;
- health examination surveys or HES - surveys which are usually a combination of questionnaires and health examination including diagnostic and laboratory tests. In HES, the contact between participants and research personnel is personal since the health examination is a component part of the survey. In this type of surveys, also questionnaires are usually communicated to the selected study subjects through personal interview.

Cross-sectional studies

There exist several similar definitions of cross-sectional studies:

- according to A Dictionary of Epidemiology (1), cross-sectional studies are studies that examine the relationship between diseases or other health-related characteristics, and other phenomena of interest in a defined population at a particular time;
- a summary of several other definitions is that cross-sectional studies are observational epidemiological studies of health status of the population in which a "snap-shot" of or a cross-section through frequency and characteristics of health outcomes and other health-related events like exposures are studied (5-9). This characteristic also gave the name to this type of epidemiological studies: cross-sectional studies are studies that measure the prevalence of health outcomes or determinants of health, or both, in a population at a specific point in time, or over a short period (7).

Health outcomes and other health-related events could be measured in cross-sectional studies on different measurement scales. In those cross-sectional studies in which the outcome event is dichotomous the prevalence of this dichotomous event is recorded. This is the reason that cross-sectional studies are also called prevalence studies (5,6, 10,11). Prevalence studies thus could be on one hand regarded as a subgroup of cross-sectional studies (11), while on the other hand all cross-sectional studies could be regarded as prevalence studies since we can dichotomize values of every observed outcome.

The selected specific point in time could be a time window within which data are collected (e.g. calendar week or month). It could also be a specific point in time in the course of events, differing in respect of each individual study subject with regard to the actual time (beginning of schooling, retirement, etc.) (6,12,13).

Frequently, cross-sectional epidemiological studies are designated as cross-sectional surveys (7).

Detailed description of cross-sectional studies' characteristics is given in numerous textbooks and handbooks including advantages and disadvantages (5,7,11,13), aims (5,6,7,9,14), methods and tools (1,4,7,10,12,15-18), and course (phases and periods) (7,10,19-21).

Although all phases/periods of cross-sectional studies' course are important, the planning period is the most important and the most sensitive period. If designing and planning the cross-sectional study in the wrong way, the whole study could be set on an inappropriate basis, and the deficiencies of this period are very difficult to be eliminated in the later phases of the cross-sectional studies. In order to avoid as many faults as possible, the course of the cross-sectional studies must be planned systematically and with all due care. A precise management of individual

cross-sectional study is very difficult to be advised. Nevertheless, common recommendations on actions in designing and planning phases of the cross-sectional studies exist (7,13,19,21,22-24).

Intervention and public health intervention

Several definitions exist of what the intervention is, among which could be find the following:

- an intervention is a generic term used to denote all public actions e.g. policies, programmes, projects (25);
- an intervention is an action or programme that aims to bring about identifiable outcomes (26).

Planned/desired effects of an intervention expressed in terms of outcomes are general objectives of an intervention.

A public health intervention is an intervention, which is applied to many, most, or all members of a community, with the aim to deliver a specific benefit to the community or population as well as benefits to individuals (26,27). Public health interventions include (26,27):

- policies of governments and non-governmental organizations;
- laws and regulations;
- organizational development;
- community development;
- education of individuals and communities;
- engineering and technical developments;
- service development and delivery; and
- communication (including social marketing).

Cross-sectional studies – an important tool in evidence-based public health

Today, cross-sectional surveys represent one of the most important tools of the evidence-based public health (14). Unfortunately, these studies are less powerful in comparison to randomized controlled trials - the main study design in evidence-based medicine. Also, the volume of evidence is smaller, and the time from intervention to outcome is longer (14). Nevertheless, they have some advantages over randomized controlled trials.

CINDI programme surveys

The World Health Organization (WHO) Countrywide Integrated Non-communicable Diseases Intervention (CINDI) programme is an intervention programme with integration as a key concept in prevention of chronic non-communicable diseases (28-30). It arose out of experiences of one of the first community-based health intervention projects in Europe - the North Karelia Project in Finland, which started in 1972 and reached remarkable achievements as well as global recognition (31).

Surveys which are aimed on one hand at assessment of burden of risk factors for non-communicable diseases, and on the other hand at evaluation of process of CINDI programme are essential components of this programme (29,32). Currently, we distinguish between two types of CINDI surveys (32,33):

- CINDI Risk Factors and Process Evaluation surveys (30,32) - this is a HES kind of surveys, and represents a basic type of CINDI surveys which provide the basic data for starting a CINDI programme in a country, and on its progress. This type of surveys is in most CINDI countries performed on a level of demonstrational area, or at most regions;
- CINDI Health Monitor Survey (33) - this is a HIS type of surveys which offer the most rough but comprehensive overview on the problems tightly associated with non-communicable diseases. This type of CINDI surveys is mostly aiming at monitoring, assessing and comparing the trends of health behaviour in CINDI countries with different politically-economic systems. Owing to comparability, monitoring should be conducted under the uniform methodology and on a national level.

Case study

Cross-sectional studies in planning public health interventions in Slovenia

CINDI programme and related surveys in Slovenia

There exist several types of surveys which provide important information for planning public health interventions for controlling non-communicable disease in Slovenia (34,35):

1. CINDI Health Monitor Survey (33-35) - This type of survey was performed in Slovenia for the first time in 2001 (36), for the second time in 2004, and for the third time in 2008. With its national and at the same time regional level, this type of surveys in Slovenia represent very strong support to development of evidence-based policy on both levels, what is extremely important in the process of diminishing interregional differences. At the same time, it is a very powerful tool for evaluation of the effectiveness of health promotion programmes. All databases include data on about 9000 participants' health behaviour.
2. CINDI Risk Factors and Process Evaluation Survey (30,34,35) - so far, there heve been three surveys performed at the demonstrational level (Ljubljana demonstrational region) – in winter 1990/1991, in winter 1996/1997, and in winter 2002/2003.

3. Ad-hoc surveys - among ad-hoc surveys, the Beltinci process evaluation surveys should be mentioned in the first place (37). With the means of two consecutive surveys of HES type, the effect of one year intervention programme in Beltinci community was evaluated. The surveys were based on CINDI Risk Factors and Process Evaluation surveys methodology.

These surveys and the resulting databases are the basis for different specific studies aiming at planning as much effective public health interventions as possible.

CINDI Health Monitor Surveys as a tool for development of effective intervention programmes for enhancing healthy nutrition and physical activity in the adult population

Background

Behavioural risk factors are among the most important risk factors for non-communicable diseases (38,39). A study showed that prevalence of some risk factors for cardiovascular diseases among which unhealthy nutrition and physical activity habits seems to be the most unfavourable ones in Eastern Slovenia (40). In order to determine population groups at highest risk for unhealthy behaviours in nutrition related to obesity and diabetes as well to determine population groups at highest risk for insufficient physical activity, a special study was performed. The intention was to prepare a high quality basis for working out the strategies, guidelines/recommendations, as well as concrete implementation action plans for long-term diminishing the high mortality attributable to non-communicable diseases (41).

Methods

The data from CINDI Health Monitor (CHM) 2001 database were used (41). The sample size was 15,379 and the age range 25-64 years. The response rate was 63.8% (9,666 respondents). The questionnaires of 9,034 respondents were eligible for analysis (eligibility criteria: sex and age provided by SORS). In the analysis of unhealthy behaviours in nutrition related to obesity and diabetes all individuals were considered, while in insufficient physical activity only 7,718 questionnaires of participants without any kind of disability (41).

Comprehensive synthetic indicators were constructed as follows (41):

- A complex indicator of unhealthy behaviours in nutrition related to obesity and diabetes was defined on the basis of several guidelines and recommendations (42-44), taking into account circumstances specific to Slovenia (cultural and economic), and possibilities of the CHM Questionnaire. The complex indicator was derived on the basis of several questions of the CHM questionnaire on nutrition habits. The containment of energy in foods was the most important criterion to select questions to be incorporated in this complex indicator. All the most important components recommended by the WHO (high/frequent intake of high energy density foods, high/frequent intake of fats, especially those composed of saturated fat acids, high/frequent intake of sweet soft drinks) (45), available in our database, were taken into consideration (41). The participants were classified into three groups on the basis of the median value on the number of unhealthy components for the whole sample as follows: healthy (0 components); moderately unhealthy (1-2 components); very unhealthy (3-7 components). The prevalence of very unhealthy behaviour was observed (41);
- A complex indicator on the average level of physical activity was derived on the basis of several questions as well. Questions were based on International Physical activity Questionnaire (46) as suggested by the CINDI WHO. There were taken into account different types of physical activity - moderate physical activity, vigorous physical activity, or walking). According to the type of physical activity and frequency (frequency of at least 4-times per week was considered as regular), participants were classified into the following groups: inactive, irregularly active, low intensity regularly active (regular walking), moderate intensity regularly active (regular moderate physical activity), and high intensity regularly active (regular vigorous physical activity). Absence of physical activity and irregular physical activity of any type or intensity were considered as insufficient physical activity and any regular physical activity (including regular walking which is one of popular types of regular physical activity in elderly in Slovenia) was considered as sufficient. The prevalence of insufficient physical activity was observed.

The observed outcomes were related to sex; age; level of education; employment; social class (self-classification); type of residence community, and geographical region.

On the basis of the logistic regression model, the risk-score for each participant was calculated and converted to the estimation of risk for the observed outcome. All participants were put in an array according to their risk estimate. Those with estimate values above the 95th percentile were classified in the very-high-risk group. The combinations of seven observed characteristics (sex, age, education, employment, social class, type of residence community, and geographical region) were then examined. Different combinations denoted different population groups' profiles. The most frequent profiles within the very-high-risk groups were observed. Those ranked as top 10 were considered as convenient for public health (PH) actions (41).

Results

Very unhealthy nutrition related to obesity

The highest odds ratios were observed in: men, aged 25-29 and 30-39 years, adults with the lowest education level (uncompleted or completed primary school), heavy workers in rural economy, people self-classified in labour social class, those living in rural communities, and those living in Eastern Slovenia.

The risk for this unhealthy behaviour was evident in 8,052 participants with data on all seven factors considered in the multivariate analysis (89.1%). The highest estimated risk score value was 0.73, while the value of the 95th percentile was 0.59. About 410 participants were classified on or above this cut-off point in the very-high-risk group for unhealthy behaviours in nutrition related to obesity and diabetes. Profiles, ranked on the top 10 ranking places according to frequency are presented in Table 1.

Table 1. Profiles, ranked on the top 10 places according to frequency within 409 respondents classified in the very-high-risk-group for very unhealthy nutrition related to obesity and diabetes: Slovenia, 2001

Profile Rank	Frequency N (%)	Risk	Sex	Age	Level of education	Employed	Social class	Residence Community	Region
1	63 (15.4)	0.62	male	30-39	vocational	yes	labour	rural	eastern
2	41 (10.0)	0.59	male	30-39	vocational	yes	middle	rural	eastern
3	30 (7.3)	0.60	male	40-49	Primary	yes	labour	rural	eastern
4.5	28 (6.8)	0.69	male	30-39	Primary	yes	labour	rural	eastern
4.5	28 (6.8)	0.62	female	30-39	Primary	yes	labour	rural	eastern
6	20 (4.9)	0.59	male	40-49	uncompleted primary	yes	labour	rural	eastern
7	18 (4.4)	0.64	male	25-29	vocational	yes	middle	rural	eastern
8	16 (3.9)	0.66	male	25-29	vocational	yes	labour	rural	eastern
9	13 (3.2)	0.59	female	25-29	vocational	yes	labour	rural	eastern
10	11 (2.7)	0.60	male	25-29	vocational	yes	labour	rural	western

Insufficient physical activity

The highest odds ratios were observed in: women, aged 25-29 years, adults with the lowest education level (uncompleted primary school), administrative/intellectual workers and job seekers, people self-classified in lowest social class, those living in urban communities, and those living in Central Slovenia.

The highest estimated risk score value was 0.38, while the value of the 95th percentile was 0.28. About 340 participants were classified on or above this cut-off point in the very-high-risk group for insufficient physical activity. Profiles, ranked on the top 11 ranking places (the profiles on 10th and 11th place had the same frequency and both had to be considered) according to frequency are presented in Table 2.

Discussion on using the survey results as a tool in planning public health interventions

With the above-described methodology we succeeded to identify population groups at highest risk for two unhealthy behaviours related to chronic non-communicable diseases:

- As for unhealthy nutritional behaviours related to obesity and diabetes the worst situation was observed in Eastern Slovenia. This was expected because, in general, the unhealthiest traditional lifestyle from the nutritional point of view was observed in the Eastern part, and the healthiest in Western Slovenia (40). This thesis was confirmed by the basic results on elements of unhealthy behaviours in nutrition related to obesity and diabetes as well.
- Healthy nutrition habits e.g. consumption of sea food and olive oil were more expressed in Western Slovenia, while unhealthy nutrition habits e.g. consumption of lard, fried food, or sweet soft drinks were most expressed in the most eastern part (47). The results indicated that from the PH point of view it was essential to start to intervene in Eastern Slovenia in terms of unhealthy nutritional behaviours related to obesity and diabetes (41).
- In insufficient physical activity there were many problems with interpretation of the results. We do strongly believe that this was the obvious consequence of the questionnaire used (41). A short last-7-days self-administered format of IPAQ is designed to observe at the same time vigorous and moderate physical activity, and walking of different sources (leisure time activities, housekeeping work, physical activity at the work-place, and transportation physical activity) (46). With regards to the impact of regular and sufficiently intensive physical activity on human health this inevitably means mixed-information data, which are less applicable for such types of observations. Despite significant amounts of energy could be spent in some of the considered types/modes of activity, not all kinds of physical activity are equally healthy - often they could be even unhealthy. Vigorous physical activity in compulsory positions of the body for a longer time periods, as it is the case in heavy physical workers in industry and rural economy, could be extremely unhealthy while

periodical vigorous physical activity during the leisure time could constitute both physical and psychological relaxation and is obviously healthy. From the viewpoint solely to the daily expenditure of energy, it is understandable that the situation with regular physical activity was shown the best in Eastern Slovenia, because the economy there is largely rural (47,48). On the contrary, another more in-depth survey, CINDI Risk Factors and Process Evaluation 2002/2003 survey, which comprised also some elements of health examination survey (and not only health interview survey), in cooperation with a multisectorial and multidisciplinary research project "Physical activity for health", showed that the prevalence of practicing regular leisure-time physical activity was the lowest in Eastern Slovenia, in rural environments, and in hard workers in rural economy (49). This indicated again that it was the priority to start with intervention programmes in the same part of the country with very similar population profiles as in nutrition, related to obesity and diabetes. On the basis of the results presented, we can conclude that in view of future research on the impact of physical activity on health, the part on physical activity in the CHM questionnaire at the international level has to be reassessed as a lot of countries have already experienced similar problems with the same set of questions (50). Also, we should try to assess the physical activity patterns of different population groups in the past, because health condition in the present is mainly influenced by physical activity and nutrition habits in the past.

Table 2. Profiles, ranked on the top 11 places according to frequency within 341 respondents classified in the "very high-risk" group for insufficient physical activity: Slovenia, 2001

Profile Rank	Frequency N (%)	Risk	Sex	Age	Level of education	Employed	Social class	Residence Community	Region
1	31 (9.1)	0.28	female	30-39	secondary	yes	middle	urban	central
2	27 (7.9)	0.32	female	30-39	university	yes	middle	urban	central
3.5	18 (5.3)	0.30	female	25-29	secondary	yes	middle	urban	central
3.5	18 (5.3)	0.29	female	30-39	university	yes	upper middle	urban	central
5	14 (4.1)	0.28	female	40-49	university	yes	upper middle	urban	central
6.5	13 (3.8)	0.31	female	40-49	university	yes	middle	urban	central
6.5	13 (3.8)	0.28	female	25-29	university	yes	middle	urban	eastern
8.5	11 (3.2)	0.34	female	25-29	university	yes	middle	urban	central
8.5	11 (3.2)	0.30	female	40-49	college	yes	middle	urban	central
10.5	10 (2.9)	0.31	female	30-39	university	yes	middle	urban	western
10.5	10 (2.9)	0.28	female	30-39	college	yes	middle	suburban	central

The results of this study, combined with results of other CINDI studies in Slovenia, confirmed that the situation of unhealthy nutritional behaviours related to obesity and diabetes, and of leisure-time physical activity, is the worst in Eastern Slovenia, especially in Pomurje. Indirectly, the situation just described could be supported also by other results of the CHM survey, which showed that the percentage of obese adults (body mass index ≥30.0) was the highest in the Pomurje (18.8% in contrast to 9.7% in most North-Western region Nova Gorica) (51).

Changing the traditional lifestyle is one of the most important elements in reducing the unhealthy behaviours of different kinds but it is extremely difficult, as the process is long lasting, and tightly bound to the political and economical situation of a country (52). In unfavourable socio-economic circumstances, the preservation of traditional lifestyle is endorsed and can be reduced only by strong multisectorial engagement (53). Such conditions currently exist in Eastern Slovenia (47,54). But, despite unfavourable socio-economic circumstances in this region which could seriously affect the success of the PH interventions, the multisectorial and multidisciplinary project "Mura", which started in 2001 in Pomurje, in only a couple of years offered several extremely positive results (37). It was a project based on intervention programmes based on the pattern of a similar process in Finland, which proved as successful and effective (31). Its development and implementation was strongly supported by the results of the presented study with its unique methodology as well by the results of related studies. Numerous multisectorial activities, including primary health prevention activities, were focused on changing the nutritional and physical activity behaviours of the population, and have been in process since the end 2001 at the regional (first in the Pomurje region) (37), as well as at the national level (55). With regard to health prevention activities, specific socio-economic and cultural circumstances were taken into consideration. On the level of population groups at-risk, the concrete health promotion and health education approach was already applied in Beltinci Community in Pomurje region (37), where the prevalence of many other unhealthy behaviours, beside insufficient leisure-time physical

activity and unhealthy behaviours in nutrition related to obesity and diabetes, is the highest (40), as well as combination of multiple risky behaviours (56). According to the first analysis of CHM survey 2004 (which serves as an efficiency evaluation tool for activities) it was shown e.g., that the prevalence of every day consumption of sweet soft drinks decreased from 42.9% in 2001 to 29.1% in 2004 (57). The same study showed strong shift to more healthy behaviour also in use of fat for food preparation. The percent of people using lard decreased from 30.3% in 2001 to 20.8% in 2004, while the percent of people using olive oil increased from 7.1% to 15.2%. Unfortunately, the comparison in physical activity behaviour was impossible since in 2004 the long last-7-days self-administered format of IPAQ was used instead of the short last-7-days self-administered format (46) in order to distinguish between physical activity from different sources (leisure time activities, housekeeping work, physical activity at the work-place, and transportation physical activity).

Conclusions

The results of this study with its unique methodology proved to be a powerful tool in development and implementation of an effective healthy nutrition and physical activity intervention programmes in Slovenia, as well as in a robust assessment of their effectiveness and efficiency. The information on the prevalence of unhealthy nutritional behaviours related to obesity and diabetes and insufficient physical activity in Slovenia, even rough, is very important for high quality health promotion and disease prevention planning at national or regional levels, since these data provide information about the comprehensive dimensions of the problem in the community.

CINDI and related surveys in developing other interventions for controlling non-communicable diseases in the adult population

The results of the above described study, as well as of the other studies based on CINDI methodology certainly serve as a basic knowledge of high-quality and applicability in preparation of PH strategies/activities in Slovenia, as well as in evaluation of their efficiency.

A similar methodology as used in analysis and identification of population at risk for unhealthy nutritional behaviours related to obesity and diabetes and insufficient physical activity was used in different other unhealthy behaviours (e.g. frequent perception of stress) and in some diseases/health states as well (e.g. musculoskeletal diseases and disorders). A detailed description on CHM Survey 2001 results is available in an extensive survey report in Slovene language for fund providers - Ministry of Education and Sports and Ministry of Health of Republic of Slovenia (58). This report is composed of several in-depth studies on different unhealthy behaviours in Slovene adults. A short version is available in the English language as well (48). Some of in-depth studies based on CHM Survey 2001 database were published in domestic or international periodicals, mostly in the English language. Additionally, interregional differences in different health phenomena inside Slovenia were possible to assess since the data enabled this kind of analyses. Chronologically these studies are as follows:

- The studies on interregional differences in health (59) and health behaviours (40);
- The study on identification of population groups at very high risk for frequent perception of stress (60) (in the English language);
- The study on identification of population groups with multiple hazardous health behaviours for cardiovascular diseases (56) (in the English language);
- The study on prevalence of selected musculoskeletal diseases and disorders in different population groups (61) (in the Slovene language);
- The study on seat-belt use and non-use in adults (62) (in the English language);
- The study on population groups at high risk for poor oral self-care (63) (in the English language), and;
- The study on self-rated health with emphasis on poor self-rated health (64) (in English language).

Also studies based on CINDI Risk Factors and Process Evaluation surveys data serve to the same purpose. They are less numerous but not less important:

- The study on efficiency of CINDI programme in controlling hypertension in the adult population of Ljubljana demonstrational level in 12-year period (65,66) (in the English language), and;
- The detailed study on prevalence of arterial hypertension, its awareness and control in the adult population of the Ljubljana area (67) (in the English language).

Some other surveys were performed in Slovenia in the last years as well. The results were published only in the Slovene language so far:

- The study on effectiveness of Beltici “Let’s Live Healthy” project (37) (in the Slovene language).
  Results of Beltinci process evaluation surveys showed considerable improvement not only in health behaviours but also in some of the physiological risk factors. The study was performed on 158 adults with monitoring/observation of health indicators on physiological risk factors before and after the intervention programme was carried out. After only one year of intervention activities, the average values of systolic blood pressure decreased by 4.7%, diastolic blood pressure by 4.1%, and blood cholesterol by 4.9%. All differences were highly statistically significant. This project was already spread from Beltinci community to other parts of Slovenia as a part of implementation of already mentioned nation-wide strategy for prevention

of cardiovascular diseases (55). The results are very promising and stimulating, but sustainability is questionable as Slovenia is still in the time of transition and the priorities are changing all the time.

- A multisectorial and multidisciplinary research project "Physical activity for health" (49) (in the Slovene language).
  This project was composed of several cross-sectional studies in the adult population, as well as in children. It was focused particularly on leisure-time physical activity. As already mentioned, results of this project showed that the prevalence of practicing regular leisure-time physical activity was the lowest in Eastern Slovenia, in rural environments, and in hard workers in rural economy.

Surveys and environmental health in Slovenia

Cross-sectional studies are most frequently used in studying health behaviour. However, they can be used to study other problems as well. In Slovenia, a study of this type was used as a lever to trigger evidence-based environmental health activities in Zasavje region which is one of most polluted regions in Slovenia (68).

Exercises

Task 1

Carefully read the part on theoretical background of this module. Critically discuss the characteristics of health surveys with your colleagues.

Task 2

From domestic (e.g. Biomedicina Slovenica, and COBISS-Cooperative Online Bibliographic System of Slovenia in Slovenia), and/or international bibliographic databases (e.g. Medline, PubMed) find out if any other health survey has been already performed in your country. If yes, then try to find out its characteristics and how its results were used in health care planning.

Task 3

If not, try to find an example from other countries (e.g. FINBALT Health Monitor Surveys).

Task 4

Discuss with your colleagues the advantages of these surveys and make proposals how they could be used more efficiently.

References

1. Last JM (editor). A dictionary of epidemiology. Oxford: Oxford University Press, 2001.
2. Anonymous. Surveys. In: Kirch W (Ed.). Encyclopedia of Public Health. New York: Springer, Science+Business Media, LLC, 2008.
3. Rossi PH, Freeman HE. Evaluation: A Systematic Approach. Newbury Park: Sage Publications, 1993.
4. Aromaa A, Koponen P, Tafforeau J (Eds.).Status and future of health surveys in the European Union. Final report of Phase 2 of the project Health Surveys in the EU: HIS and HIS/HES Evaluations and Models. Helsinki, Bruxelles: National Health Institute KTL, Scientific Institute of Public Health; 2003.
5. Abramson JH. Cross-sectional studies. In: Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press, 2002. p.509-528.
6. Hennekens CH, Buring JE. Epidemiology in medicine. Boston, Little, Bown and Company, 1987.
7. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC, 1999.
8. Beaglehole R, Bonita R, Kjellstrom. Basic epidemiology. Geneva: World Health Organization, 1998.
9. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia, W.B. Saunders company, 2001.
10. Premik M. Introduction to epidemiology (in Slovene). Ljubljana: Faculty of Medicine, 1998.
11. Pearce N. A Short Introduction to Epidemiology. Second Edition. Wellington, New Zealand: Centre for Public Health Research, Massey University Wellington Campus, 2005.
12. Bailey KD. Methods of social research. New York: The Free Press (A Division of Macmillan, Inc.), 1987.
13. Vaughan JP, Morrow RH. Manual of epidemiology for district health management. Geneva: WHO, 1989.
14. Brownson RC, Baker EA, Leet T, Gillespie KN. Evidence based public health. Oxford, New York: Oxford University Press, 2003.
15. Hansen MH, Hurwitz WN, Madow WG. Sample survey methods and theory. Volume I Methods and applications. New York: John Wiley Sons, 1993.
16. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research. New York: Van Nostrand Reinhold, 1982.

17. Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press, 2002.
18. Altman DG. Practical statistics for medical research. Chapman & Hall: 1993.
19. Lutz W, Chalmers J, Lockerbie L, Hepburn W. Health and community surveys. Volume 1. A practical manual for health and community workers. Geneva: IEA, World Health Organization, MacMillan, 1992.
20. Lutz W, Chalmers J, Lockerbie L, Hepburn W. Health and community surveys. Volume 2. A practical manual for health and community workers. Geneva: IEA, World Health Organization, MacMillan, 1992.
21. Aday LA. Designing and conducting health surveys: a comprehensive guide. San Francisco: Jossey-Bass, 1996.
22. Bankowski Z, Bryant JH, Last JM (editors). Ethics and epidemiology: international guidelines. Proceedings of the XXVth CIOMS Conference. Geneva: Council for International Organizations of Medical Sciences, 1991.
23. World Medical Association. World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. Available at: URL:
www.who.int/bulletin/archives/79(4)373.pdf. Accessed: June 4, 2013.
24. Public Health Leadership Society. Principles of the ethical practice of public health. New Orleans: Public Health Leadership Society; 2002. Available at: URL:
http://209.9.235.208/CMSuploads/PHLSethicsbrochure-40103.pdf. Accessed: June 4, 2013.
25. Planning and Performance Management Unit. Office of the Director of U.S. Foreign Assistance. Glossary of Evaluation Terms. Washington D.C.: Planning and Performance Management Unit. Office of the Director of U.S. Foreign Assistance, 2009.
26. Rychetnik L, Hawe P, Waters E, Barratt A, Frommer M. A glossary for evidence based public health. J Epidemiol Community Health 2004;58;538-545.
27. McLeroy K, Norton B, Kegler M, et al. Community-based intervention. Am J Public Health 2003;93:529-33.
28. World Health Organization, Regional Office for Europe. A strategy to prevent chronic disease in Europe. A focus on public health action. The CINDI vision. Copenhagen: WHO, Regional Office for Europe; 2004.
29. World Health Organization, Regional Office for Europe. CINDI Countrywide Integrated Non-communicable Diseases Intervention Programme. Protocol and guidelines for monitoring and evaluation procedures. Berlin: Springer-Verlag; 1987.
30. World Health Organization, Regional Office for Europe. Handbook for process evaluation in noncommunicable disease prevention. CINDI 2000. Copenhagen: World Health Organization, Regional Office for Europe, 1999.
31. Puska P, Tuomilehto J. Nissinen A, Vartainen E. The North Karelia project. 20 year results and experiences. Helsinki, National Public Health Institute; 2001.
32. WHO, Regional Office for Europe. CINDI 2000. Protocol and guidelines. Copenhagen: WHO, 1996.
33. Prattala R, Helasoja V, Laaksonen M, Laatikainen T, Nikander P, Puska P. Cindi health monitor. Proposal for practical guidelines. Helsinki: National Public Health Institute, 2001.
34. Zaletel-Kragelj L, Fras Z, Maucec Zakotnik J. Monitoring of the situation regarding cardiovascular diseases in adult Slovene population and its role for quality assurance in their treatment (in Slovene). Slov. Kardiol 2006;3:4-12.
35. Maucec Zakotnik J, Fras Z, Zaletel-Kragelj L. The WHO Countrywide Integrated Non-communicable Diseases (CINDI) programme in Slovenia. In: Donev D, Pavleković G, Zaletel-Kragelj L (Eds.). Health promotion and disease prevention: a handbook for teachers, researchers, health professionals and decison makers. Lage: Hans Jacobs publishing company, 2007. p. 204-219. Available at: URL:
http://www.snz.hr/ph-see/publications.htm. Accessed: June 4, 2013.
36. Zaletel-Kragelj L. Risk factors for noncommunicable diseases in adults in Slovenia Ljubljana. Survey protocol. (in Slovene) Ljubljana: Institute of Public Health, Faculty of Medicine and CINDI-Slovenia, 2001.
37. Zelko-Peterka E, Zaletel-Kragelj L. Local community health enhancing programme assessment. The case of Beltinci community (in Slovene). Slov Kardiol 2005;2:8-12.
38. National Institute of Public Health Determinants of the burden of disease in the European Union. National Institute of Public Health, Stockholm, 1997.

39. World Health Organization (2002) The world health report 2002: reducing risks, promoting healthy life. World Health Organization, Geneva.
40. Zaletel-Kragelj L, Erzen I, Fras Z. Interregional differences in health in Slovenia: II. Estimated prevalence of selected behavioral risk factors for cardiovascular and related disease. Croat Med J 2004;45:644-50. Dostopno na: URL: http://www.cmj.hr/2004/45/5/15495295.pdf. Accessed: June 4, 2013.
41. Zaletel-Kragelj L, Fras Z, Maucec Zakotnik J. Results of CINDI Health Monitor surveys in Slovenia as a tool for development of effective healthy nutrition and physical activity intervention programmes. J Public Health 2006, 14: 110-118. Available from: URL: http://www.springerlink.com/media/927ywwwtml6x73yyrqau/contributions/x/1/2/2/x122x18g238p15v7.pdf.Accessed: June 4, 2013.
42. European Commission. Status report on the European Commission's work in the field of nutrition in Europe. EC, Bruxelles, 2002.
43. World Health Organization. Recommendations for preventing diabetes. In: World Health Organization. Diet, nutrition and the prevention of chronic diseases. Report of a joint WHO/FAO expert consultation. WHO (WHO Technical Report Series 916), Geneva, 2003. p. 72-80.
44. World Health Organization. Integrated prevention of noncommunicable diseases. Draft global strategy on diet, physical activity and health. WHO (WHO EB113/44), Geneva, 2003.
45. World Health Organization. Diet, nutrition and the prevention of chronic diseases. Report of a joint WHO/FAO expert consultation. World Health Organization (WHO Technical Report Series 916), Geneva, 2003.
46. IPAQ Group. International Physical Activity Questionnaire. Available from: URL: https://sites.google.com/site/theipaq/. Accessed: June 4, 2013.
47. Hanzek M (editor). Human development report. Slovenia 1999. Institute of Macroeconomic Analyses and Development & United Nations Development Programme, Ljubljana, 1999.
48. Zaletel-Kragelj L, Fras Z, Maucec Zakotnik J. Health behaviour and health among Slovene adult population, 2001 CINDI Health Monitor Survey 2001. University of Ljubljana, Faculty of medicine, 2005.
49. Zavrsnik J, Pisot R, Fras Z, Zaletel-Kragelj L, Strel J, Sila B, Zerjal I, Zurc J, Kropej In: Physical activity for health: presentation of results of Targeted Research Programme "Competitive position of Slovenia 2001-2006" (in Slovene). Univerza na Primorskem, Znanstveno-raziskovalno sredisce, Institut za kinezioloske raziskave, Koper, 2004.
50. World Health Organization, Regional Office for Europe. CINDI Health Monitor: A Study of feasibility of a health behaviour monitoring survey across CINDI countries. Data book. WHO, Regional Office for Europe, Copenhagen, 2003.
51. Zaletel-Kragelj L, Erzen I. Assessment of selected health conditions and selected risk factors. In: Zaletel-Kragelj L, Erzen I, Belović B, Artnik B, Premik M, Zelko-Peterka E. Assessment of interregional differences in health in Slovenia and searching for solutions for diminishing them. Final report on project of Targeted Research Programme (in Slovene). Univerza v Ljubljani, Medicinska fakulteta, Katedra za javno zdravje, Ljubljana, 2003. p 15-86.
52. Wilson AJ, Oldenburg BF, Lopez AD. Targeted approaches for reducing inequities in chronic disease. MJA 2003;179:231-232.
53. Oldenburg B. Preventing chronic disease and improving health: broadening the scope of behavioural medicine research and practice. Int J Behav Med 2002;9:1-16.
54. Javornik J, Korosec V (editors.). Human development report. Slovenia 2003. Institute of Macroeconomic Analyses and Development & United Nations Development Programme, Ljubljana, 2003.
55. Fras Z, Mauzcec Zakotnik J (2002) Life-style interventions in primary health care – implementation of Slovenian CINDI approach as a national strategy for prevention of coronary heart disease. In: International conference: Promoting health through physical activity and nutrition. Radenci, Slovenia, April 18-21, 2002. Final programme: abstracts and papers. CINDI Slovenija, Ljubljana, pp 87-93.
56. Zaletel-Kragelj L, Fras Z. Identification of population groups with multiple hazardous health behaviours for cardiovascular diseases in Slovenia. Slov Kardiol 2005;2:112-117.
57. Zaletel-Kragelj L, Mauzcec Zakotnik J, Belović B. Assessment of health promotion activities in Health region Pomurje. First results of "CINDI Health Monitor Slovenia 2004" (in Slovene). In: Fras Z (ed.) Slovene Forum for cardiovascular diseases prevention 2005: book of articles. Zdruzenje kardiologov Slovenije, Ljubljana, pp 27-32.

58. Zaletel-Kragelj L, Fras Z, Mauzcec Zakotnik J (editors). Risky behaviours related to health and selected health conditions in adult population of Slovenia: results of Slovenia CINDI Health Monitor Survey 2001 (in Slovene). Ljubljana: CINDI Slovenia, 2004.
59. Zaletel-Kragelj L, Erzen I, Fras Z. Interregional differences in health in Slovenia: I. Estimated prevalence of selected cardiovascular and related diseases. Croat Med J 2004;45:637-643. www.cmj.hr/2004/45/5/15495294.pdf. Accessed: June 4, 2013.
60. Zaletel-Kragelj L, Pahor M, Bilban M. Identification of population groups at very high risk for frequent perception of stress in Slovenia. Croat Med J 2005;46:153-161. Available from URL:
www.cmj.hr/2005/46/1/15726688.pdf. Accessed: June 4, 2013.
61. Djomba JK. Estimated prevalence of selected musculoskeletal diseases and disorders in different population groups in Slovene adults (in Slovene). Department of Public Health, Faculty of Medicine, University of Ljubljana, Ljubljana; 2005.
62. Bilban M, Zaletel-Kragelj L. Seat-belt use and non-use in adults in Slovenia. Int J Public Health 2007;52:317–325.
63. Artnik B, Zaletel-Kragelj L. Population groups at high risk for poor oral self-care: the basis for oral health promotion. Int J Public Health 2008;53:195-203.
64. Farkas Lainszcak J, Zaletel Kragelj L. Self-rated health in Slovenian adults. Slov Kardiol 2008;5:42-49.
65. Bulc M, Fras Z, Zaletel-Kragelj L. Twelve-year Blood Pressure Dynamics in Adults in Ljubljana Area, Slovenia: Contribution of WHO Countrywide Integrated Non-communicable Diseases Intervention Program. Croat Med J 2006;47:469-77. Available from: URL: http://www.cmj.hr/2006/47/3/16758526.pdf. Accessed: June 4, 2013.
66. Zaletel-Kragelj L, Mauzcec Zakotnik J, Fras Z. Public health intervention programmes and their evaluation. In: Donev D, Pavleković G, Zaletel-Kragelj L (editors). Health promotion and disease prevention: a handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs publishing company, 2007. p. 145-155. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: June 4, 2013.
67. Fras Z, Bulc M, Maucec Zakotnik J, Zaletel-Kragelj L. Prevalence of arterial hypertension, its awareness and control in the adult population of the Ljubljana area of Slovenia. Results of WHO's countrywide integrated noncommunicable diseases intervention (CINDI) program survey 2002/2003. Slov Kardiol, 2006;3:106-114.
68. Kukec A, Farkas-Lainszcak J, Erzen I, Zaletel-Kragelj L. A prevalence study on outdoor air pollution and respiratory diseases in children in Zasavje, Slovenia, as a lever to trigger evidence-based environmental health activities. Arh Hig Rada Toksikol. 2013;64:9-22.

Recommended readings

1. Abramson JH. Cross-sectional studies. In: Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press, 2002. p.509-528.
2. Cross-sectional surveys. In: Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC, 1999. p.213-229.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Cluster analysis**
Module: 2.15	ECTS (suggested): 1.0
Author(s), degrees, institution(s)	Anca Vitcu, PhD, Assistant Professor Faculty of Computer Science, University "Al. I. Cuza", Iasi, Romania
Address for correspondence	Anca Vitcu Faculty of Computer Science, University "Al. I. Cuza", Iasi Str. General Berthelot, nr. 16, Iasi, Code 700483, Romania Fax: +40 232 201490 Email: avitcu@yahoo.com
Keywords	Hierarchical clustering, k-means clustering, probabilistic clustering.
Learning objectives	After completing this module students and public health professionals should: • understand and apply the basic clustering methods; • be aware of cluster analysis limitations; • gain expertise in the interpretation of clustering solutions; • improve the skills needed for the use of statistical software packages; • be able to undertake investigations into geographical patterns of disease; • identifie and investigate disease risks in certain areas; • increase knowledge regarding the efficient distribution of resources for prevention and treatment of diseases.
Abstract	From the huge amount of data available in the health field today, we have to find out in a very short time based on developed computer science technologies how an area is doing, what the emerging problems are and recommend the best scenario we can get to solve them. In light of this, we have to learn to use the most appropriate tools, which can help us to mine the "mountains" of data around us with the aim to create knowledge. The chapter below is a brief presentation of clustering analysis which includes a number of different algorithms and methods accustomed to organize a huge amount of observed data into meaningful structures.
Teaching methods	Teaching methods will include combination of lectures, exercises, individual work, interactive methods such as small group discussions. Before the introductory lecture, a case study could be presented to increase students' motivation. An introductory lecture gives the students the basic theoretical knowledge on cluster analysis. The theoretical approach should be illustrated with exercises and study cases from the field of epidemiology. After the introductory lecture students will work individually and in teams of 2-3 people, study the recommended readings and discuss the characteristics and pitfalls of clustering algorithms. The work will be followed by an individual case problem presentation and overall discussion.
Specific recommendations for teachers	• ECTS: 1.0; • work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a computer room; • equipment: computers (one computer for 2-3 students), LCD projection equipment, internet connection, access to the bibliographic data-bases; • training materials: recommended readings and software packages or other related documents; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment could be based on case problem presentations.

CLUSTER ANALYSIS

Anca Vitcu

Theroretical background

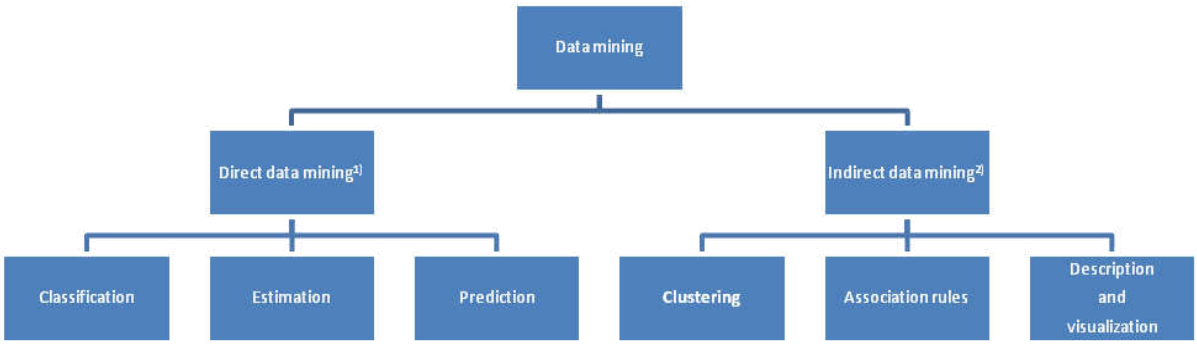
About clustering

Definition and main features

Cluster analysis (clustering), segmentation analysis, taxonomy analysis, or unsupervised classification, is a data analysis tool which comprises a variety of goals, all of them related with grouping or segmenting the data (objects, instances, items) into meaningful structures (clusters), without explaining why they exist, based only on the data that describes them and their relationships.

Clustering represents an important component of data mining, which is a process of exploring and analyzing large amounts of data in order to discover valuable information (i.e. some patterns, relationships among all the variables) (1):

Diagram 1. Data-mining methods



It includes a number of different algorithms and methods and is used when we do not have any a priori hypothesis about the data. The components of a cluster are more closely related to one another than those assigned to different clusters. In other words, the degree of association between objects is maximal if they belong to the same cluster and minimal, otherwise.

A relevant problem concerns the evaluation of clustering quality. The quality criterion should be dependent on the aims of the clustering and should be designed considering that the results of the process should meet the research requirements.

Aims

The main objective of cluster analysis is to identify homogeneous groups or clusters in a data set. Cluster analysis has the potential to generate new knowledge: may help formulate hypotheses concerning the origin of the sample, describe a sample in terms of a typology, or predict the future behaviour of population types. It also proved to be a useful starting point for other analysis procedures (e.g. based on geographical analysis methods and spatial scan statistics localised clustering can be assessed and the risk of disease inside and outside the study area can be compared) (2,3).

Applications [10]

In medical sciences, there are different types of clustering such as:

- General clustering – involves the analysis of the overall clustering tendency of the disease incidence in a study area without searching the exact location of the clusters (3,4)
- Specific clustering – involves specific disease clustering methods which are designed to examine the exact location of the clusters (5,6)

Some of the most important applications of cluster analysis are: pattern recognition, spatial data analysis, image processing, gene expressions.

[1] in direct data mining some variables are selected as targets.
[2] in indirect data mining no variables are selected as targets – we are not sure what we are looking for, what plays a role in forming some relationships, or how these relationships do that.

In epidemiology, clustering is frequently used:

- To identify diseases and their stages, and by examining their characteristics to discover whether there are different subtypes of diseases grouped togheter under a single diagnosis (7).
- To detect patterns in the spatial and temporal distribution of a disease (8).
- To find if there are several distinct groups of patients with different symptoms and similar behavior habits which have been diagnosed with a certain disease (9).
- To group patients into non-overlapping activity/inactivity clusters and then use the outcomes in models of prevalent and incident overweight.
- To measure the different effects of treatments on classes within the population.
- To assist the disease surveillance.

Cluster analysis for gene expression data can be performed in two ways:

- To group genes with similar expression patterns.
- To group different samples on the basis of corresponding expression profiles.

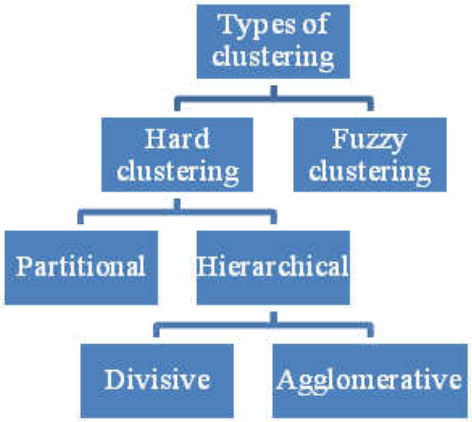
Description

Types of clustering

In a cluster analysis we can distinguish different types of clustering (10,11):
(A) exclusive (hard) or non-exclusive (fuzzy);
(B) complete or partial;
(C) hierarchical or partitional.

In the following paragraphes, we will shortly describe each of these clustering types. A relationship among the concepts is presented in Diagram 2 (1):

Diagram 2. Categories of clustering algorithms



(A) Exclusive vs. non-exclusive clusters

Exclusive clusters assign each case to a single cluster so as each case is closer to all of the cases in its cluster than to any case in another cluster (Figure 1). In this situation, a case can only belong to one cluster (e.g. case 1 belongs to cluster “c”, case 2 to cluster “a”, case 3 to cluster “c”, case 4 to cluster “a” and so on).

On the other hand, there are situations in which a case can logically be placed in more than one cluster, e.g. case 2 belongs to cluster “a” and “c”, case 4 belongs to cluster “b” and “c” (Figure 2). These situations in which a case can simultaneously belong to more than one cluster or is “located between” two or more clusters, and can be assigned to any of them are addressed by non-exclusive clustering.

An example of non-exclusive clustering is fuzzy clustering, where every case belongs to every cluster with a membership weight which takes values between 0 (absolutely doesn’t belong) and 1 (absolutely belongs). In fuzzy clustering it is often added the constraint that the sum of weights for each case must equal 1. In a similar way, probabilistic clustering techniques compute the probability with which a case belongs to each cluster, and these probabilities must also sum to 1. These approaches are most appropriate for avoiding the arbitrariness of assigning a case to only one cluster when logically it may belong to several. In practice, a fuzzy or probabilistic clustering is often converted to an exclusive clustering by assigning each case to the cluster in which its membership weight or probability is the highest (12-15).

Figure 1. Exclusive clusters

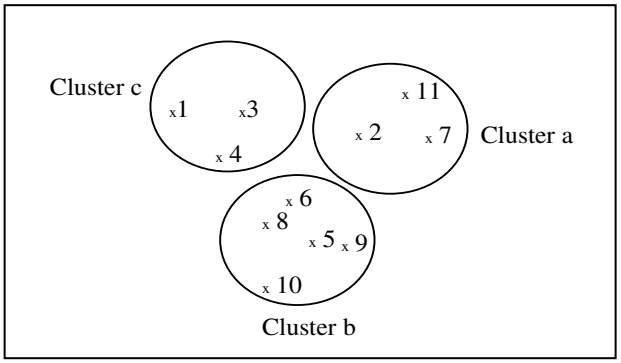
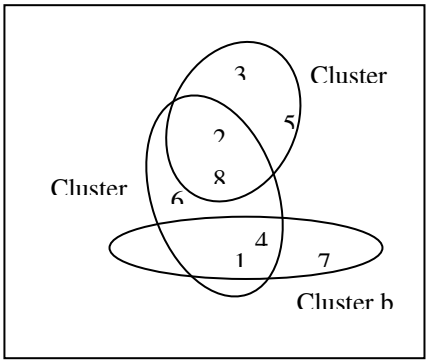


Figure 2. Non-exclusive clusters



A clustering of a given data set D can be represented by a function defined as follows:

$f : D \to [0,1]^k,\ \mathbf{x} \to f(\mathbf{x})$,

where $\mathbf{x} \in D$, and $f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \cdots\cdots \\ f_k(\mathbf{x}) \end{pmatrix}$,

with $f_i(\mathbf{x}) \in [0,1]$ for $i = 1, 2, \ldots, k$, and $\sum_{i=1}^{k} f_i(x) = 1 \quad \forall \mathbf{x} \in D$.

If for every $\mathbf{x} \in D$, $f_i(\mathbf{x}) \in \{0,1\}$, then the clustering represented by f is a hard clustering; or else, it is a fuzzy clustering (1).

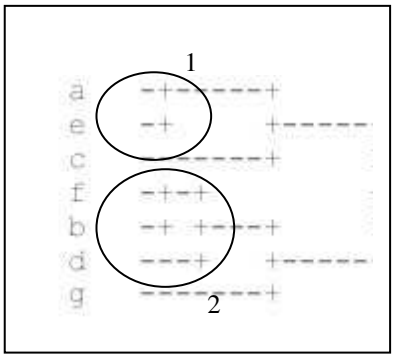
(B) Complete vs. partial clusters

A complete clustering assigns every case of a data set to a cluster while a partial clustering is used when some cases in a data set may not belong to well-defined groups (e.g. they may represent outliers, or uninteresting issues).

(C) Hierarchical vs. partitional clusters

The goal of a cluster analysis is sometimes to arrange the cases or clusters themselves into a natural hierarchy. This involves successively grouping them so that at each level of the hierarchy, clusters (resp. cases) within the same group (resp. cluster) are more similar to each other than those in different groups (resp. clusters).

Figure 3. Hierarchical clusters



In figure 3, we can see that the two subclusters with components (a,e) and (c) belong to cluster 1 while the other two subclusters composed of cases (f,b,d) and (g) respectively belong to cluster 2. In other words, if a cluster has subclusters we obtain a hierarchical clustering. We will say more about hierarchical clustering later in this chapter.

The opposite of hierarchical clustering is partitional clustering which is a division of the set of data into exclusive clusters.

Fundamental steps in cluster analysis

In principal, the basic steps in a clustering analysis are the following (10,14):

1. Selection of the appropriate variables on which cluster configuration will be based.

Note: The initial choice of variables determines the characteristics that can be used to identify clusters; if important variables are excluded, poor or misleading findings may result.

2. Selection of the appropriate cases to be included in the analysis.

Note: The clusters' structure could be quite different among selected groups: women vs. men, socioeconomic groups, life style, regions or countries.

3. Selection of the appropriate measures for calculation the distance or similarity between objects.

Note: Distance is a measure of how different two objects are, while similarity is a measure of their closeness. Objects which are alike share a low distance and a high similarity. Selection of a distance/similarity measure should be based on both the properties of the measure and on the algorithm for cluster formation.

4. Selection of the appropriate methods to be used for combining objects into clusters.

Note: Selection of a method depends on the data set to be analyzed, definition of measure supplied to it and the characteristics of the various methods available. Different methods will result in different cluster patterns.

5. Application of these methods: there are various clustering algorithms that can be used to analyse a given data set; it is important to choose a good clustering algorithm for the given data set, and to evaluate appropriately the selected clustering method.

6. Evaluation of the results.

Note: Basically this refers to:

- set up the number of clusters (if it is not known),
- test the clusters stability,
- test the validity of clusters (internal, relative and external validity).

7. Interpretation of the results: for a succesful analysis, clusters need to have a substantive interpretation.

In the following sections we will give a brief description of the main features of the steps mentioned above.

Assumptions regarding variables

In cluster analysis variables should be commensurable, which means that they must have equal scales. Frequently, in a data set we have to deal with the following situations:

- The variables are quantitative but have different scales (e.g. age, income, human development index)
- The variables included in the analysis are mixed, in other words they have different measurement levels: quantitative (e.g. age, weight, blood pressure), ordinal (e.g. attitude scale: 1=strongly agree, 2=agree, 3=disagree, 4=strongly disagree), nominal (e.g. types of health services, development areas)
- The occurrence of one variable depends on another variable (e.g. occupational level depends on occupational status)

If variables are quantitative or ordinal and have different scales, a transformation to an equal scale is required. The most popular transformations are (14):

1. z **-** standardization [(a) theoretical or (b) empirical] – transforms the values so that they have a mean of 0 and a standard deviation of 1:

a. $z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}$, where x_{ij} are values of case i in variable j, μ_j is the mean value derived from the attributes of the scale and σ_j is the standard deviation derived from the attributes of the scale;

b. $z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$, where x_{ij} are values of case i in variable j, $\bar{x}_j$ is the empirical mean of variable j and s_j the empirical standard deviation of variable j.

2. [0,1] – transformation [(a) theoretical or (b) empirical] - the procedure subtracts the minimum value from each item being standardized and then divides by the range:

a. $z_{ij} = \frac{x_{ij} - \alpha_j}{\beta_j - \alpha_j}$, where x_{ij} are values of case i in variable j, α_j is the theoretical minimum of the scale and β_j is the theoretical maximum of the scale;

b. $z_{ij} = \frac{x_{ij} - a_j}{b_j - a_j}$, where x_{ij} are values of case i in variable j, a_j is the empirical minimum of the scale and b_j is the empirical maximum of the scale.

Thus, if the variables are quantitative or ordinal, the types of scales we have to deal with take account of: continuous values, and resp. discret equidistant values or discret non-equidistant values. For each of these scales we can compute the variables' theoretical mean and standard deviation according to the subsequent formula:

- $\mu_j = (\beta_j + a_j)/2$ and $\sigma_j = (\beta_j + \alpha_j)/2\sqrt{3}$ - for continuous values
- $\mu_j = (\beta_j + a_j)/2$ and $\sigma_j = \sqrt{(n_j - 1)(n_j + 1)/12}$, where n_j is the number of categories of variable j - for discret equidistant values
- $\mu_j = (\sum v_{jk})/n_j$ and $\sigma_j = \sum (v_{jk} - \mu_j)^2 / n_j$, where n_j is the number of categories of variable j and v_{jk} scale value of category k of variable j - for discret non-equidistant values

The following alternatives are also available for transforming values:

- Range -1 to 1 - each value for the item being standardized is divided by the range of the values;
- Maximum magnitude of 1 - the procedure divides each value for the item being standardized by the maximum of the values;
- Mean of 1 - the procedure divides each value for the item being standardized by the mean of the values;
- Standard deviation of 1 - the procedure divides each value for the variable or case being standardized by the standard deviation of the values.

Note: Standardization can be done by variables or by case.

Another situation mentioned above concerns the case of mixed variables. In this circumstance the following transformations can be approched:

- $x'_{ij} = x_{ij}$ for binary variables,
- $x'_{ij}(n) = \begin{cases} 1/\sqrt{2} & \text{if } x_{ij} = n \\ 0 & \text{otherwise} \end{cases}$ for nominal variables
- $x'_{ij} = x_{ij}/r$ or $x'_{ij} = (x_{ij} - \min(x_j))/r$, where r is the range for ordinal and quantitative variables

The resulting variables have values between 0 and 1.

If transformation $1/r$ causes problems, a priory standardization of the variables can overcome them.

Assumptions regarding data - Missing values

In almost all databases we can find observations which have missing values in one or more of the variables. The most common method of including missing values in dissimilarity calculation is to omit each observation pair having at least one missing value (listwise deletion), or a case is eliminated only if the number of missing values exceeds a certain threshold (pairwise deletion) (16). These methods can fail in the circumstance when both observations have no measured values in common.

An alternative is to estimate missing values with imputation techniques. Both observations could be imputed using the mean or median of each variable over the non-missing data.

For categorical variables, the value "missing" can be considered as another categorical value, if both objects have missing values on the same variables.

Distance and proximity measures

Fundamental to all clustering techniques is the choice of distance or dissimilarity measure between two objects.

An appropriate way to introduce a measure is by means of a metric function, $d: X \times X \to R^1$ which satisfies the following conditions for all $x_i, x_{i'}, x_{i''} \in X$:

1. $d(x_i, x_{i'}) \geq 0$ (non-negativity);
2. $d(x_i, x_{i'}) = 0$ if and only if $x_i = x_{i'}$;
3. $d(x_i, x_{i'}) = d(x_{i'}, x_i)$ (symmetry);
4. $d(x_i, x_{i'}) \leq d(x_i, x_{i''}) + d(x_{i''}, x_{i'})$ (triangle inequality).

Generally, in cluster analysis a metric is considered only as a framework for dissimilarities. In many applications not all of the above conditions are required.

We first discuss distance measures before describing a variety of algorithms for clustering.

There are various measures that can be used to quantify similarity or dissimilarity between objects (15). These measures can be classified into four main groups:

1. correlation coefficients,
2. distance measures,

3. derived measured based on correlation coefficients or distances,
4. other similarity or dissimilarity measures developed for special purposes, mainly for binary variables.

Note: Correlation coefficients and derived measures based on correlation coefficients are mostly used for clustering variables while, distance measures and derived measures based on distances are mostly used for clustering cases.

In most cases in cluster analysis we work with raw data, but sometimes we don't have them. However, the data available is characterized in terms of proximity between pairs of cases (objects).

Example 1: patients from the same type of hospitals located in ten different counties are asked to judge by how much certain medical services differ from one another. The patients are selected randomly from the hospital database and asked to indicate their attitude toward each of the 7 items that refers to the quality of health services. The attitude is measured on a seven point Likert scale with the following response categories: 1=very favorable, 2=moderately favorable, 3=slightly favorable, 4=neither favorable nor unfavorable, 5=slightly unfavorable, 6=moderately unfavorable, 7=very unfavorable. Dissimilarities can then be computed by averaging over the collection of such judgments. The results can be represented by a symmetric matrix $\left(d_{ij}\right)_{n\times n}$ with non-negative entries and zero diagonal elements, where n represent the number of objects and d_{ij} the proximity between cases i^{th} and j^{th}. This matrix is then provided as input to clustering algorithm. Table 1 envisages the proximity matrix of the imaginary situation presented above. Together with this cluster analysis certain statistics are also recommended to be computed: mean, percent of favorable, percent of unfavorable, percent of neutral, standard deviation, etc.

Table 1. Proximity matrix of patients attitude: values are average pairwise dissimilarities of counties from a questionnaire given to sampled patients

	County1	County2	County3	County4	County5	County6	County7	County8	County9	County10
County1	0									
County2	5.27	0								
County3	6.34	7.12	0							
County4	3.42	3.82	4.92	0						
County5	2.25	4.67	6.17	3.67	0					
County6	6.17	6.92	4.50	5.67	4.23	0				
County7	4.95	3.57	3.92	5.00	5.25	3.92	0			
County8	6.33	2.67	4.81	2.24	3.00	4.52	5.83	0		
County9	4.75	4.51	6.83	4.25	8.12	6.00	2.85	3.77	0	
County10	7.00	5.00	6.08	3.02	4.58	5.02	5.76	6.42	6.92	0

In most of the studies, data are measurements for n cases on k variables. As we know, the popular clustering algorithms require the specification of a dissimilarity matrix as their input. For this reason, before deciding the type of clustering algorithm, we have to construct the dissimilarity matrix:

$$D\left(x_i, x_{i'}\right) = \sum_{j=1}^{k} d_j\left(x_{ij}, x_{i'j}\right),$$

where $d_j\left(x_{ij}, x_{i'j}\right)$ is the distance between values of j^{th} variable for cases i and i'. These measures have to be calculated according to the measurement level of variables: quantitative, ordinal, or categorical.

Quantitative variables

First, we would like to remember that measurements of this type of variables are represented by continuous real-valued numbers. Among the distance measures the following ones are the most popular:

Euclidian distance: $d\left(x_i, x_{i'}\right) = \sqrt{\sum_j \left(x_{ij} - x_{i'j}\right)^2}$

Squared Euclidian distance: $d\left(x_i, x_{i'}\right) = \sum_j \left(x_{ij} - x_{i'j}\right)^2$ - places more weight on larger differences than smaller ones.

Note: (1) Euclidian distances are usually computed from raw data and not from standardized ones (2). They can be affected by differences in scale among the dimensions from which the distances are computed.

City-block (Manhattan) distance: $d\left(x_i, x_{i'}\right) = \sum_j \left|x_{ij} - x_{i'j}\right|$

Chebychev distance: $d(x_i, x_{i'}) = \max_j |x_{ij} - x_{i'j}|$

Minkowski distance: $d(x_i, x_{i'}) = (\sum_j |x_{ij} - x_{i'j}|^p)^{1/p}$, where $p \geq 1$.

Note: For p = 2, p = 1, and p = ∞ the Euclidian, the Manhattan and the Chebychev distances are, respectively, obtained.

Alternatively, for these types of variables clustering can be based on correlation measures.

Pearson's r can be used as a similarity or correlation measure:

$\rho(x_i, x_{i'}) = \frac{\sum_j (x_{ij} - \bar{x}_i)(x_{i'j} - \bar{x}_{i'})}{\sqrt{\sum_j (x_{ij} - \bar{x}_i)^2 \sum_j (x_{i'j} - \bar{x}_{i'})^2}}$, where x_{ij} is the value of case i in variable j, $x_{i'j}$ the value of case i' in variable j, $\bar{x}_i = \sum_j x_{ij} / p$, $\bar{x}_{i'} = \sum_j x_{i'j} / p$.

If inputs are first standardized, then we can write $\sum_j (x_{ij} - x_{i'j})^2 = 2(1 - \rho(x_i, x_{i'}))$, which imply that clustering based on correlation is equivalent to that based on squared distance.

Ordinal variables

The values of this type of variables are represented as adjacent integers, and the possible values are considered to be an ordered set (e.g. degrees of preferences or agreement). Rank data are special kind of ordinal data.

The most common measures for ordinal data are: city block metric, coefficient kappa for ordinal variables, correlation coefficients (Kendal's tau or Gamma).

Categorical variables

In this case, the degree of difference between pairs of values must be delineated explicitly.

For clustering variables the following measures can be used: Cramer's V, Phi, Lambda and other association coefficients.

If cases are clustered the following measures can be applied for nominal variables: simple matching coefficient, coefficient kappa for nominal variables, city block metric, squared Euclidian distances.

Binary variables

The measures for binary variables differ in the importance they attach to the different cells of a 2x2 table as the one presented below:

		Case i	
		Present (1 or +)	Absent (0 or -)
Case i'	Present (1 or +)	a	b
	Absent (0 or -)	c	d

Reading the table, we find out that for "a" variables both cases have value "present", for "d" variables both cases have value "absent", for "b" and "c" variables the cases have different values.

There are situations when we may want to weight the positive-positive cell more or less than the negative-negative cell, we may want to weight equal mismatches and matches cells or we may want to ignore one of them, etc. according to the information they provide. In this context, for binary variables the following measures can be used (14,17):

Jaccard's coefficient I: $\rho(x_i, x_{i'}) = d/(2d + b + c)$

Dice's coefficient: $\rho(x_i, x_{i'}) = 2d/(2d + b + c)$

Sokal&Sneath's coefficient I: $\rho(x_i, x_{i'}) = d/(d + 2(b + c))$

Russel&Rao's coefficient: $\rho(x_i, x_{i'}) = d/(a + b + c + d)$

Euclidian distance: $d(x_i, x_{i'}) = \sqrt{b + c}$ - has a minimum value of 0 and no upper limit

Squared Euclidian distance: $d(x_i, x_{i'}) = b + c$ - has a minimum value of 0 and no upper limit

Lance and Williams (non-metric dissimilarity measure): $d(x_i, x_{i'}) = (b + c)/(2a + b + c)$ – has a range of 0 to 1

Pattern distance: $d(x_i, x_{i'}) = bc/(a + b + c + d)^2$ - has a range of 0 to 1

Mixed levels

Various methods have been developed to work with mixed measurement levels. One of these methods is based on Gower's dissimilarity coefficient:

$$d_{ii'} = \frac{\sum_{i'} w_{ii'j} \cdot d_{ii'j}}{\sum_{i'} w_{ii'j}}$$

or Gower's similarity coefficient

$$s_{ii'} = \frac{\sum_{i'} w_{ii'j} \cdot s_{ii'j}}{\sum_{i'} w_{ii'j}},$$

where $d_{ii'}$ and $s_{ii'}$ are the dissimilarity and respectively similarity between cases i and i', $w_{ii'j}$ is the weight for variable j, $d_{ii'j}$ and $s_{ii'j}$ are the dissimilarity and respectively similarity between cases i and i' in variable j.

While squared Euclidian distance and city block metric can be computed for all measurement levels they are often used as dissimilarity measures for Gower's coefficient. In each situation the weight is defined as the inverse value of the maximum distance such as its highest value is 1.

Clustering methods

Different clustering methods with different properties have been developed. A general answer to the question "which technique should be used?", cannot be given. The answer depends on the data used and the analysed question.

An important distinction is the question whether "cases or variables should be clustered?" and in this situation an answer can be provided.

Broadly speaking, we can identify three different general types of cluster analysis algorithms (18):

- those based on a hierarchical attempt to discover cluster structure,
- those based on an attempt to find the optimal partition into a specified number of clusters,
- those based on a probabilistic model for the underlying clusters.

In the following paragraphes we will give a brief discription to each of these algorithms.

Different methods have been developed to cluster cases or variables which result in two major types of assignments (19,20):

- **Deterministic**
  - Hierarchical algorithms (divisive hierarchical algorithms, agglomerative hierarchical algorithms) – according to this algorithm we construct step by step a hierarchy or tree-like structure to see the relationship among objects,
  - Non-hierarchical algorithms (K-means algorithms) – consistent with this method a position in the measurement is taken as central location, and distance between cases is measured from such central point (21).
- **Probabilistic**
  - Probabilistic algorithms – according to this method there is a probability or a degree of membership with which the case belongs to each of the cluster, such as in the example presented in Table 2 (e.g. case 1 belongs with a probability of 0.2 to cluster 1, with a probability of 0.3 to cluster 2 and with a probability of 0.5 to cluster 3, we will see later in this chapter that ussualy, exclusive clusters are required and for this reason the case is assigned to the cluster which has the highest probability).

Table 2. Deterministic vs. probabilistic assignment of objects in cluster analysis

Objects	Cluster membership	Deterministic assignment			Probabilistic assignment		
		Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
1	3	0	0	1	0.2	0.3	**0.5**
2	3	0	0	1	0.3	0.0	**0.7**
3	1	1	0	0	**0.8**	0.2	0.0
4	2	0	1	0	0.3	**0.5**	0.2
5	1	1	0	0	**0.7**	0.1	0.2
6	2	0	1	0	0.2	**0.6**	0.2
7	2	0	1	0	0.0	**0.9**	0.1

Example 2: Periodically, the management team of hospital X conducts interview surveys with persons who benefit of the hospital assistance. The respondents are patients selected randomly from the

registration list. As part of the survey each patient is asked to indicate her/his perception toward a battery of Likert-type items regarding the quality and safety of health services provided by the hospital (staff behavior and attitude, hospital facilities, etc.). If the management team wants to analyze what services are alike, they have to use methods for clustering variables (e.g. agglomerative hierarchical algorithms). If they want to find out if patients differ in their attitude and dissimilar patterns of preference can be recognized, cases have to be clustered. Theoretically, all three algorithms mentioned above (hierarchical, non-hierarchical, and probabilistic) can be used for this purpose.

Note: (1) Hierarchical techniques may be used to cluster cases or variables while K-means and probabilistic methods only to cluster cases. (2) Hierarchical algorithms can be applied for small and moderate sample sizes. K-means algorithms require at least a moderate sample size while probabilistic algorithms require large sample size.

Hierarchical algorithms

Hierarchical algorithms involve a concept of ordering motivated by the number of observations that can be combined at a time or the assumption that the distance between two observations or clusters is not statistically different from 0.

Hierarchical algorithms can be classified into divisive algorithms and agglomerative algorithms (17,22,23).

Divisive methods start with the assumption that all objects are part of a single cluster. The algorithm splits this large cluster step by step until each object is a separate cluster.

Agglomerative methods start inversely, are bottom-up procedures, and ended when all observations are combined in one cluster. Most common agglomerative methods are (Diagram 3):

- **Single linkage** (nearest neighbor approach) – The method works in the following way: At the first step, each cluster consists of one object. At the next step, we agglomerate those two observations that have the shortest distance. A third observation, which has the next least distance, is added to the two observation cluster to create a third observation cluster or is combined with another observation to form a two observation cluster. The clusters are combined step by step. In each step, those two clusters with the smallest dissimilarity or the highest similarity are merged. Iteration continues until all objects are in one single cluster. Single linkage leads to chaining and may result in too few large and heterogeneous clusters.
- **Complete linkage** (furthest neighbor approach) – The method is similar to single linkage except that this is based on maximum distance. This result in a very strong definition of the homogeneity of clusters: The largest dissimilarity between all objects of one cluster should be less than a certain value. The farthest neighbour of each object should have a distance less than a certain value. Complete linkage results in dilatation and may produce too many clusters

  Note: Complete linkage and single linkage are extreme procedures with completely different properties.
- **Average linkage within groups** (within-groups linkage) is the mean distance between all possible inter- or intra-cluster pairs. – The method is based on the fact that the average distance between all pairs in the resulting cluster is made to be as small as possibile. This method is therefore appropriate when the research purpose is homogeneity within clusters. This method try to avoid the effects mentioned above (chaining and dilatation).
- **Centroid method** – In this situation the distance between two clusters is determined as the difference between centroids. Cluster to be merged is the one with the smallest sum of Euclidean distances between cluster means for all variables.
- **Median method** (incremental sum of squares method) – This method is similar with the centroid method but included weighting to control the differences in clusters sizes. It also uses Euclidean distance as the proximity measure.
- ***Ward's method*** – This one is a method distinct from all other methods presented above because it uses an analysis of variance approach to evaluate the distance between clusters (it attempts to minimize the sum of squares of any two clusters).

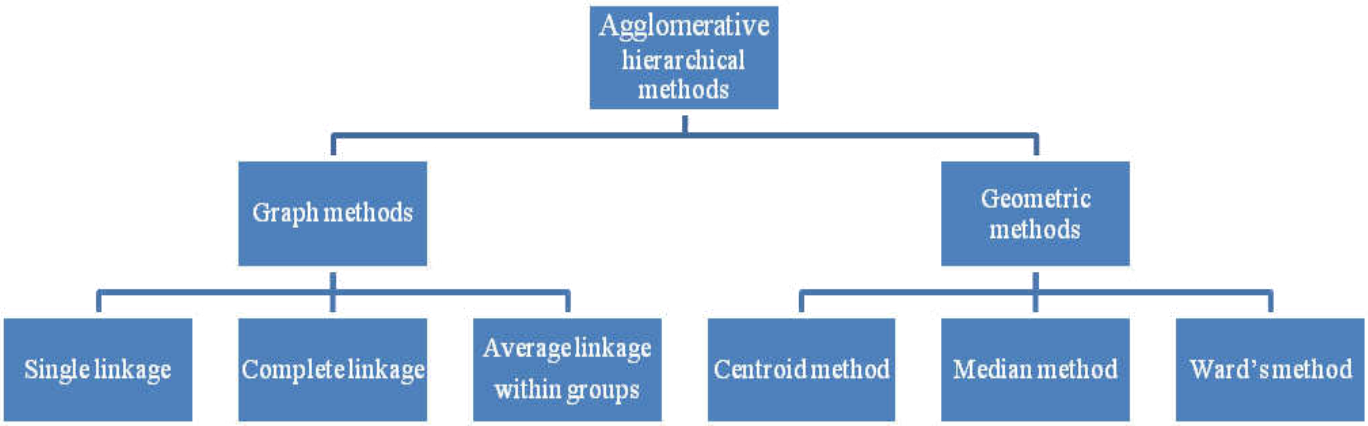
All these methods differ in the way similarities or dissimilarities are re-computed after two clusters are merged. The last three methods (centroid, median and Ward) have a different approach based on two important assumptions:

1. there is a data file to work with (all other methods require only a dissimilarity or similarity matrix that can be computed from a data file but that can be observed directly too);
2. clusters can be described by their centres (means in the variables).

In these methods the centres are computed step by step based on minimization or maximization of a certain criteria. Ward's method minimises the within sum of squares, the centroid method and median method select in each step those clusters whose centres are closest. They differ in the way the centres are calculated. The methods are primarily designed for clustering cases. The methods require squared Euclidean distances, in other words are applied to interval-scaled variables or variables that can be treated as interval-

scaled and the distances are weighted implicitly. A larger distance in one variable has a higher weight than small distances in many variables.

Diagram 3. Some common agglomerative hierarchical methods (1)



Evaluation of the results

The answer to the question "how many clusters we found?" is quite difficult regardless the clustering method applied. None of the methods usually results in a unique solution.

In most of the cases the number of clusters is determined on the basis of a dendrogram (called also hierarchical tree diagram) by counting the number of clusters that combine objects at a convenient distance level (14).

Another frequently used method is the scree test. In this case a graphic is constructed such as the x-axis contains the number of clusters, and the y-axis the agglomeration levels. A sharp increase in the agglomeration schedule results in an elbow knick.

Another important issue regarding the evaluation of the outcomes of a cluster analysis concerns the stability of a cluster solution. Different methods can provide different clusters. A cluster solution is said to be stable, if a small modification of the method specified and the data used does not change the results too much. A stability analysis of methods is usualy based on modifications of clustering techniques and dissimilarity (resp. similarity) measure (24).

Example 3: A common interest in epidemiological practice is to look for dissimilarities among geographical or economical regions regarding a certain illness (2,26). Based on the data provided by WHO database we apply hierarchical cluster analysis to group countries for which standardized death rate on diabetes is available. The data we are working with are from 2005, for males of 0-64 ages. In this framework we apply three of the methods discussed above (Ward method, single linkage and complete linkage). Squared Euclidian distance is used as dissimilarity measure. The outcomes are presented in Figures 4, 5 and 6.

Analysing the dendograms we see that all three methods involved provide the same number of clusters (seven) but moderatly different solutions.

The problem became more complex if beside standardized death rate on diabetes we include in the analysis other variables such as: total health expenditure as % of gross domestic product (GDP), health at current prices (% of total household consumption expenditure). In this situation standardization is requested (25).

According to Figure 4, using Ward method the solution is composed of the following seven clusters:

Cluster I: Greece, UK, Luxembourg, Ireland, Belarus, Iceland, Spain, Russian Federation

Cluster II: Latvia, Romania, Norway, Czech Republic, Finland, Ukraine, Slovenia

Cluster III: Kazakhstan, Lithuania, Slovakia, Poland, Croatia

Cluster IV: Malta, Estonia, Republic of Moldova

Cluster V: Austria, Kyrgyzstan, Serbia

Cluster VI: Hungary

Cluster VII: Tajikistan, Uzbekistan

Figure 4. Hierarchical cluster analysis based on Ward method

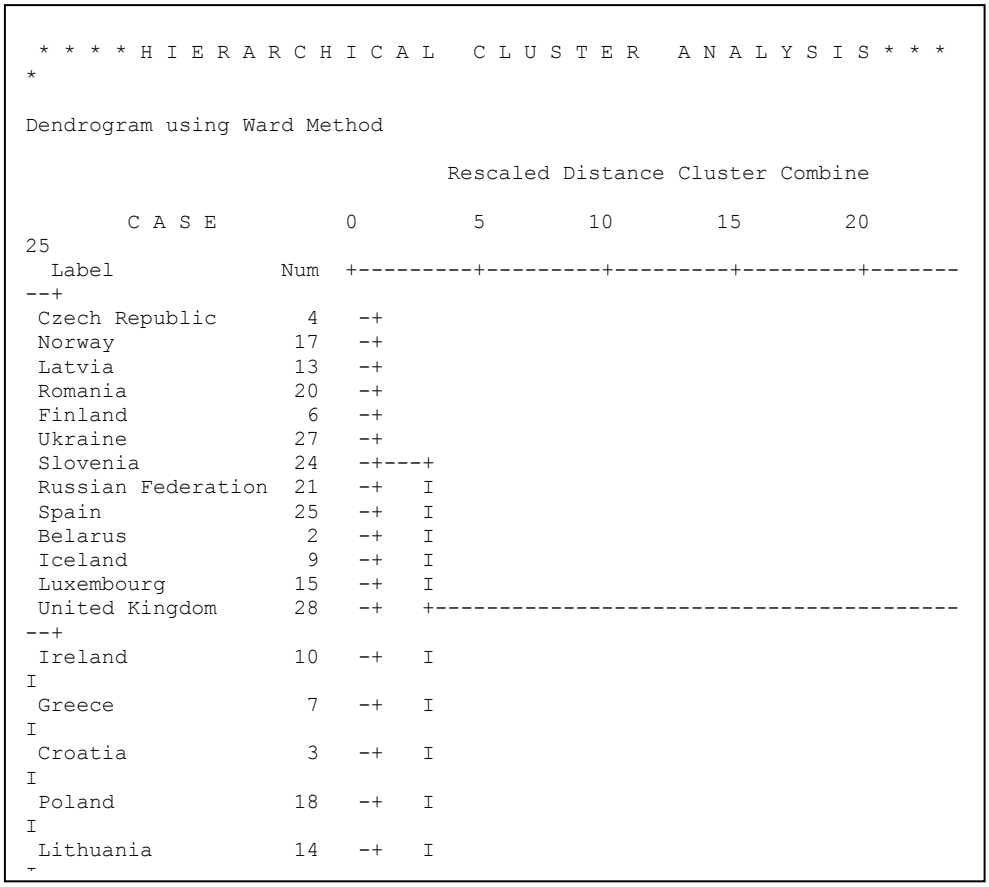
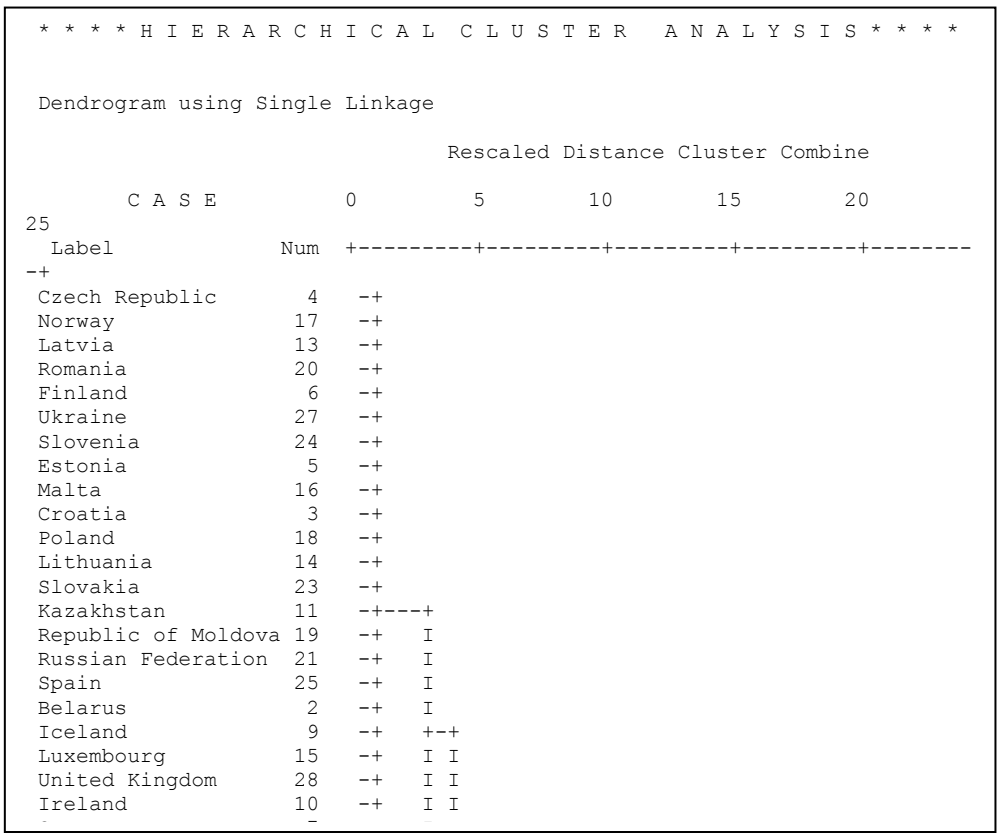


Figure 5. Hierarchical cluster analysis based on single linkage method



Analyzing the Figure 5 it can be seen that the single linkage method provided the folowing seven clusters solution:

Cluster I: Greece, UK, Luxembourg, Ireland, Belarus, Iceland, Spain, Russian Federation
Cluster II: Latvia, Romania, Norway, Czech Republic, Finland, Ukraine, Slovenia, Kazakhstan, Lithuania, Slovakia, Poland, Croatia, Malta, Estonia
Cluster III: Republic of Moldova
Cluster IV: Austria, Kyrgyzstan, Serbia
Cluster V: Hungary
Cluster VI: Tajikistan
Cluster VII:Uzbekistan

Complete linkage method provided the folowing seven clusters solution:

Cluster I: Greece, UK, Luxembourg, Ireland, Belarus, Iceland, Spain, Russian Federation
Cluster II: Latvia, Romania, Norway, Czech Republic, Finland, Ukraine, Slovenia, Kazakhstan, Lithuania, Slovakia, Poland, Croatia
Cluster III: Malta, Estonia, Republic of Moldova
Cluster IV: Austria, Kyrgyzstan, Serbia
Cluster V: Hungary
Cluster VI: Tajikistan
Cluster VII:Uzbekistan

Non-hierarchical algorithms: k-means clustering

The k-means clustering is a top-down procedure which belongs to a more general group of clustering techniques known as partitioning or optimization methods.

The job in this type of clustering is to partition a data set into k disjoint sets of objects such that the objects within each set are as homogeneous as possible. Homogeneity is captured by an appropriate score function such as minimizing the distance between each object and the centroid (average) of the cluster to which it is assigned. Often the centroid belonging to a cluster is considered to be a representative point for that cluster, and there is no explicit statement of what sort of shape of cluster is being sought. Maximizing (or minimizing) the score function is a computationally search problem. Iterative algorithms based on local search are particularly common for cluster analysis (30).

The general idea in this type of clustering is to start with a randomly chosen clustering of the objects, then to reassign objects so as to give the greatest increase (or decrease) in the score function, then to recalculate the updated cluster centers, to reassign points again, and so forth until there is no change in the score function or in the cluster memberships.

This approach has the advantage of being simple and guaranteeing at least a local maximum (minimum) of the score function. The major drawback of the search algorithms is that we do not know how good the clustering that it converges to is relative to the best possible clustering of the data (the global optimum for the score function being used).

The classical iterative algorithm used to find the optimal partition is called k –means clustering, which has close connection to the EM algorithm, about which we will talk later. In this case, we have to fix in advance the number of clusters we require (this is a typical problem of many clustering algorithms). The number of clusters is represented by the parameter k.

There are several variants of the k -means algorithm (21). The first step of the basic version involves the chosen at random of k objects to represent initial cluster centers. In the next step all objects are assigned to the nearest cluster center according to Euclidean distance, the mean value of the objects in each cluster is computed to form its new cluster center, and iteration continues until there are no changes in the clusters.

When the number of clusters cannot be specified in advance, we can apply an incremental clustering method based on a hierarchical grouping of objects which use a measure of cluster quality or, a statistical clustering method based on a mixture model of different probability distributions, one for each cluster (31,32).

Suppose we are using k -means but do not know the number of clusters in advance. A solution is to test different possibilities and see which is best, which one minimizes the total squared distance of all objects to their cluster center. In this context, a possible strategy is to start from a given minimum, k= 1, and work up to a small fixed maximum, using cross-validation to find the best value. Another strategy is to begin by finding few clusters and determining whether it is worth splitting them. Suppose k=2, perform k -means clustering until it terminates, and then consider splitting each cluster.

One way to split a cluster is to make two new seeds: a seed one standard deviation away from the clusters center in the direction of its greatest variation, and the other seed the same distance in the opposite direction. Next, apply k -means to the points in the cluster with these two new seeds. If the split can be retained, try splitting each cluster further. The process continues until no splits remain.

Note: The standard k-means algorithm can be divided into two phases: 1. the initialization phase - the algorithm randomly assigns the cases into k clusters; 2. the iteration phase - the algorithm computes the distance between each case and each cluster and assigns the case to the nearest cluster.

The standard k-means algorithm. (1)

input: data set D, number of clusters k, dimensions d:
{ C_i is the i^{th} cluster}
{1. Initialization phase}
1. $(C_1, C_2, \ldots, C_k) =$ Initial partition of D.
{2. Iteration phase}
2. repeat
3. d_{ij} = distance between case i and cluster j ;
4. $n_i = \arg\min_{1 \le j \le k} d_{ij}$;
5. assign case i to cluster n_i ;
6. recompute the cluster means of any changed clusters above;
7. until no further changes of cluster membership occur in a complete iteration
8. output results.

Partition-based methods of cluster analysis begin with a specified number of clusters and search through possible allocations of objects to clusters to find an allocation that optimizes some clustering score function. A large variety of score functions can be used to determine the quality of clustering and a wide range of algorithms has been developed to search for a good partition.

In the following section we will present some basic features of score functions.

Score functions

In order to define the clustering score function we need to look at within cluster variation and between cluster variation of a clustering C.

The within cluster variation measures how compact the clusters are, while the between cluster variation looks at the distances between different clusters.

Suppose that we have selected cluster centers, noted r_k, from each cluster. This can be a designated representative data point (object) that is defined to be "central" in some manner. If the input objects belong to a space where means have sense, we can use the centroid of the objects in the cluster C_k as the cluster center and r_k, will be defined by formula:

$$r_k = \frac{1}{n_k} \sum_{x \in C_k} x ,$$

where n_k is the number of objects in the k^{th} cluster.

Within-cluster variation can be defined as the sum of square of distances from each point to the center of the cluster:

$$within\_c(C) = \sum_{k=1}^{K} within\_c(C_k) = \sum_{k=1}^{K} \sum_{x(i) \in C_k} d(x, r_k)^2 ,$$

when $d(x, r_k)$ is defined as Euclidian distance, $within\_c(C)$ is referred to as the within-cluster sum-of-squares.

Between-clusters variation can be defined by the distance between cluster centers:

$$between\_c(C) = \sum_{1 \le j < k \le K} d(r_j, r_k)^2 .$$

In this framework, the score function of a clustering C can then be defined as a monotone combination of the factors $within\_c(C)$ and $between\_c(C)$, such as the ratio $between\_c(C) / within\_c(C)$.

The k - means algorithm, uses the means within each group as cluster centers and Euclidean distance for d to search for the clustering that minimizes the within cluster variation.

If we are given a candidate clustering, it is important to know how difficult is it to evaluate $within\_c(C)$ and $between\_c(C)$. Computing $within\_c(C)$ takes O(n) operations, while $between\_c(C)$ can be computed in $O(k^2)$ operations. Hence, computing a score function for a single clustering requires a pass through the whole data.

After we get the clusters based on the algortithm described above we have to proceed with a stability analysis. In k-means clustering the idea of stability analysis is to check, whether modifications of methods or data have a negative effect on the results.

In contrast to hierarchical methods, the technique and the distance measure (squared Euclidean distances) are fixed for k-means, in other words they cannot be modified. In this case, the only thing which is not fixed is the starting partition. Therefore, the stability of the results can be tested by modifying the starting partitions. Having in view this aim, the following strategies can be applied:

- Generate different random starting partitions, if random starting values are used.
- Re-order the cases.
- Change the starting values, if centers are entered or computed using a hierarchical technique.
- Use different starting procedures (e.g. randomly generated starting values).
- If a classification is stable, the starting procedure should have no influence.

Regarding the data stability, both cases and variables can be analyzed. In this framework cases stability can be verified considering the subsequent steps:

- Divide the data set in M sub-datasets.
- Run for each sub-dataset k-means and save the cluster centres.
- Compare the cluster centres.

M is usually set equal to 2. This method only allows comparing cluster centers. If classifications are to be compared, the strategy has to be modified in the following way:

- Divide the population in M (usually M = 2) subpopulations.
- Use one subpopulation as reference population and compute the cluster centres for this reference population.
- Run two cluster analysis for the other subpopulation: an 'ordinary' k-means analysis and a 'confirmatory' k- means analysis with fixed centres. Use the centres of the solution of the previous step.
- Compute an index to compare the classifications within each subpupolation analysed above (e.g. Rand index).

Variables stability can be tested by adding randomly distributed variables.

Instead of testing stability attempts have been made to change the algorithm in order to find a more robust classification, a best solution or to compute a partition of partitions. Experience proves that a more robust classification can be obtained using the following methods (33,34):

- Eliminating outliers in a first stage (the outliers may be assigned to clusters in the next stage).
- Using the city block metric instead of squared Euclidean distances, because it is less sensitive towards outliers.
- Weighting variables automatically according to their contribution to separate the clusters.

In this way, variables with a high proportion of random noise should be eliminated from analysis.

Note: k-means cluster analysis is very sensitive to outliers and is recommended to remove them before starting the analysis.

Two-step clustering

Two-step clustering is a method preferred for large data sets and when categorical variables with three or more levels are involved. The algorithm follows the subsequent steps:

1. Pre-clusters are identified.
2. The pre-clusters identified in the first step are treated as single cases and clustered hierarchical.

Note: k-means cluster analysis and two-stage cluster analysis usually generate different solutions.

Probabilistic clustering

In the case of probabilistic clustering the problems of k-means clustering are avoided:

- The problem of incommensurability does not occur – variables of different measurement levels and different scale units can be analysed without any transformation of the variables.
- Each case is assigned probabilistic to a cluster — usually to the cluster from which it is most likely to have come.
- The model has a statistical basis (the typically score function is likelihood of the data).

From a probabilistic perspective the goal of clustering is to find the most likely set of clusters given the data. The foundation for statistical clustering is a statistical model called finite mixture.

A mixture is a set of k probability distributions, representing k clusters that govern the variables values for members of those clusters.

Each distribution gives the probability that a particular case would have a certain set of variables values if it were known to be a member of that cluster. Each cluster has a different distribution. Any particular case belongs to one and only one of the clusters, but it is not known which one. In the end the clusters are not equally likely (35).

The simplest finite mixture situation occurs when there is only one numeric variable, which has a normal distribution for each cluster, with different means and variances. The clustering problem is to take a set of objects and a specified number of clusters and work out each cluster mean and variance and the population distribution between the clusters. The mixture model combines several normal distributions and its probability density function looks like a "mountain range with a peak for each component" (32).

Suppose we have three clusters A, B and C and each has a normal distribution with means and standard deviations: μ_A and σ_A for cluster A, μ_B and σ_B for cluster B, μ_C and σ_C for cluster C. Samples are taken from these normal distributions using cluster A with probability p_A, cluster B with probability p_B and, cluster C with probability p_C such as $p_A + p_B + p_C = 1$. We also suppose that we have a set of objects and want to determine the parameters that characterize the model. If we knew which of the three distributions each object (case) came from, the finding of the parameters is easy, the only thing we have to do is to estimate the mean and standard deviation for the three clusters separately, using the clasical formulas: $\mu = \frac{x_1 + x_2 + \ldots + x_n}{n}$, for the mean and $\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_n - \mu)^2}{n-1}$, for the standard deviation.

If we knew the parameters finding the probabilities that a given object comes from each distribution would be easy. Given an object x the probability that it belongs to cluster A is given by formula:

$$\Pr[A|x] = \frac{\Pr[x|A] \cdot \Pr[A]}{\Pr[x]}.$$

The final result of the method is not a particular cluster but rather the probabilities with which case x belongs to cluster A, B and C.

Note: the k-means algorithm has some important properties: 1) works on numerical data; 2) is efficient in clustering large data sets; 3) does not work effectively on high-dimensional data; 4) usually stops at a local optimum; 5) the designed clusters have convex shapes; 6) the performance depends on the initialization of the centers.

It is important to reveal that the standard k-means algorithm has several variations developed to improve its performance: the k-harmonic algorithm, the fuzzy k-means algorithm, and the Gaussian EM algorithm.

The EM (Expectation Maximization) Algorithm

The EM algorithm is defined as "a general statistical method of maximum likelihood estimation in the presence of incomplete data that can be used for the purpose of clustering".

The general objective of the EM algorithm is to discover clusters in observations (or variables) and to assign those observations to the clusters. An application for this type of analysis is marketing segmentation - a marketing research study in which a number of consumer behaviors related variables are measured for a large sample of respondents. The purpose of the study is to detect groups of customers that are similar to each other when compared to respondents that "belong to" other clusters. In addition to identifying such clusters, it is also interesting to determine how the clusters are different, to determine the specific variables or dimensions that vary and how they vary in regard to members in different clusters (35,36).

Among the important features of the EM algorithm we can mention: simplicity, stability, and robustness to noise.

While the k-means clustering refers to the fact that given a fixed number of k clusters observations are assigned to those clusters so that the means across clusters (for all variables) are as different from each other as possible, the EM algorithm extends this approach to clustering in two significant ways:

1. Instead of assigning cases or observations to clusters to maximize the differences in means for continuous variables, the EM clustering algorithm computes probabilities of cluster memberships based on one or more probability distributions. In this situation the goal of the clustering algorithm is to maximize the overall probability of the data, given the (final) clusters.
2. Unlike the classic implementation of k-means clustering, the general EM algorithm can be applied to both continuous and categorical variables.

The basic approach and logic of this clustering method is as follows: suppose we measure a single continuous variable in a large sample of observations and that the sample consists of two clusters of observations with different means and standard deviations. The goal of EM clustering is to estimate the means and standard deviations for each cluster so as to maximize the likelihood of the observed data.

The results of EM clustering are different from those computed by k-means clustering. The latter will assign observations to clusters to maximize the distances between clusters. The EM algorithm does not compute actual assignments of observations to clusters, but classification probabilities.

Finding the right number of clusters in k-Means and EM clustering: v-fold cross-validation

An important question that needs to be answered before applying the k-means or EM clustering algorithms is how many clusters there are in the data. This is not known a priori and there is no unique answer.

An estimate of k can be obtained from the data using the method of cross-validation. Remember that the k-means and EM methods will determine cluster solutions for a particular user-defined number of clusters. The k-means and EM clustering techniques can be optimized and enhanced for typical applications in data mining (15).

To determine k we can use v-fold cross-validation algorithm for automatically determining the number of clusters in the data (37).

This algorithm is useful in all general "pattern-recognition" tasks - to determine the number of market segments in a marketing research study, the number of distinct spending patterns in studies of consumer behavior, the number of clusters of different medical symptoms, the number of weather patterns in meteorological research, etc.

The general idea of the method is to divide the overall sample into a number of v folds. The same type of analysis is then successively applied to the observations belonging to the training sample, and the results of the analyses are applied to sample v to compute some index of predictive validity. The results for the v replications are aggregated to yield a single measure of the stability of the respective model.

Note: the EM algorithm proceeds iteratively; each iteration involves two steps: 1. the expectation step (E-step); 2. the maximization step (M-step).

Evaluation of clustering algorithms

An essential topic which should also be considered refers to the assessment of clusters validity. To compare the clustering results of different clustering algorithms, it is necessary to develop some validity criteria.

There are three fundamental criteria to investigate the cluster validity: internal criteria, external criteria and relative criteria.

The assumptions that has to be checked in the analysis of internal validity are:

- The clusters should be homogenous.
- The clusters should be different in structure.
- The classification should be able to explain the variation in the data.

In the case of external validity it has to be analysed if the clusters correlate with the external variables that are known to be correlated with the classification and that are not used for clustering.

The analysis of the relative validity implies the examination of the following constraints (27,28):

- The classification should be better than the null model that assumes no clusters are present.
- The classification should be better than other classifications.

Diagram 4 includes a presentation of some popular validity index criteria. A criteria has good or bad results in terms of the framework they are employed: some of them work well when the clusters are compact, but not when the clusters have arbitrary shape.

Internal and external criteria involve statistical testing, while relative criteria does not. Few comments about the diagram 4 are mentioned in the next caption (1).

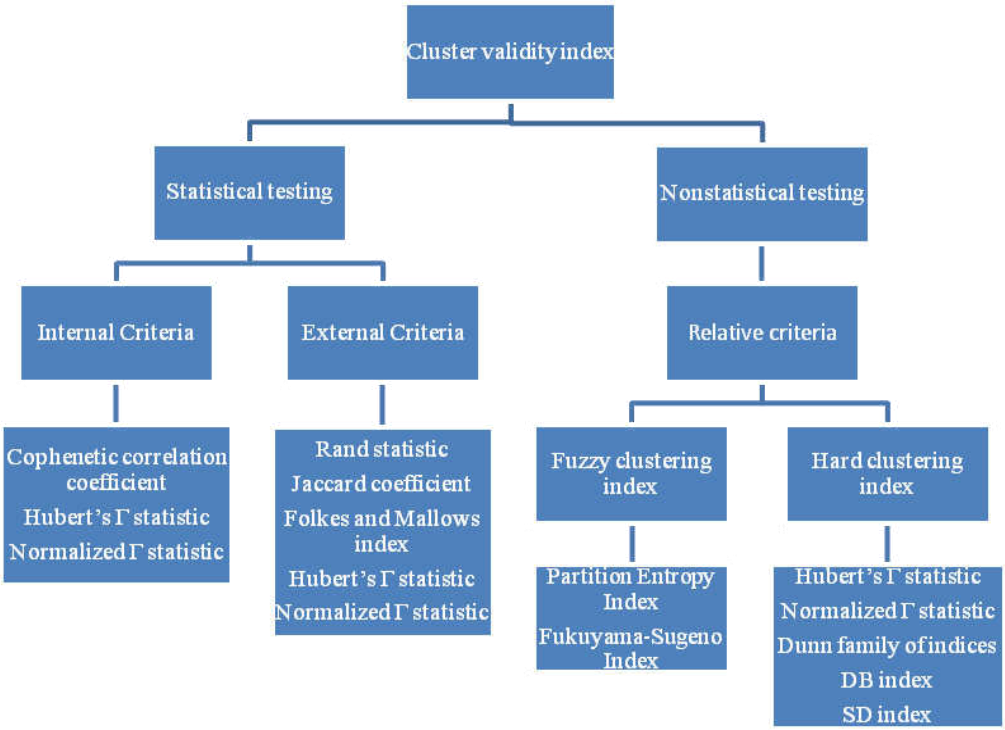
Statistical testing

Hypothesis Testing: the fundamental idea is to test if the data set has a random structure (H_0). There are proposed the following forms for the randomness hypotheses: 1. the random graph hypothesis; 2. the random label hypothesis; 3. the random position hypothesis.

Note: Monte Carlo and bootstrapping are computer simulation tools recommended in statistical testing of hypotheses.

Internal Criteria: evaluates the clustering structure produced by an algorithm using only quantities and features inherited from the data set. The method can be employed for: 1. hierarchy clustering schemes applying Cophenetic correlation coefficient; 2. single clustering schemes applying Hubert's Γ statistic or Normalized Γ statistic.

Diagram 4. Some cluster validity indices



External Criteria: the outcomes of the clustering algorithm is evaluated based on an intuitive given structure of the data set. There are two different approaches:

1. The resulting clustering structure is compared with an independent partition of the data built on the knowledge and intuition about the clustering structure of the data set.

Notation: $C = \{C_1, C_2, \ldots, C_s\}$ - the resulting clustering structure; $P = \{P_1, P_2, \ldots, P_r\}$ - an independent partition of the data; a - the number of pairs of data points which are in the same cluster of C and in the same cluster of P; b - the number of pairs of data points which are in the same cluster of C but in different clusters of P; c - the number of pairs of data points which are in different clusters of C but in the same cluster of P; d - the number of pairs of data points which are in different clusters of C and in different clusters of P; M - the total number of pairs of data points in the data set.

Then we have the following formulas:

$$M = a + b + c + d = \frac{n(n-1)}{2},$$

where n is the number of data points in the data set.

Rand statistic: $R = \frac{a+d}{M}$, with the range in [0, 1];

Jaccard coefficient: $J = \frac{a}{a+b+c}$, with the range in [0, 1];

Folkes and Mallows index: $FM = \sqrt{\frac{a}{a+b} \cdot \frac{a}{a+c}}$, with the range in [0, 1].

High values of the above mentioned indices indicate great similarity between the resulting clustering structure and the independent partition of the data.

2. The proximity matrix is compared with an independent partition of the data built on the knowledge and intuition about the clustering structure of the data set. The Hurbert's Γ statistic ([0,1]), or Normalized Γ statistic ([-1, 1]) is calculated.

Non-statistical testing

Relative criteria: the fundamental idea is to decide on the best clustering outcome from a set of given schemes according to a predefined criterion. The validity indices are grouped in two categories dependent on the clusters status, whether they are disjoint or not – if a data point (a case) belongs to only one cluster, or to two or many.

The last step but not less important in clustering analysis refers to the clusters interpretability (29). To have a clear and valid picture on the results, the outcomes have to be discussed within interdisciplinary teams.

Software products

Through a large variety of software packages available for developing cluster analysis, the most popular are the following: R packages, MATLAB programs, SAS, SPSS, CLUSTAN, ALMO, WEKA, WinBUGS.

Exercises

Task 1

What role does variable and data selection play in cluster analysis? Please give a solid argumentation of your answer. Provide examples.

Task 2

Is standardization needed in cluster analysis? Why? Please give a solid argumentation of your answer. Provide examples.

Task 3

How can we use clustering methods to identifying the outliers in a dataset?

Task 4

What happens if in cluster analysis, distance measure is replaced by similarity measure?

Task 5

What is the most important problem with non-hierarchical algorithms?

Task 6

List the factors that affect the stability of clustering solution. Select three of them and propose a method that drives to the right solution.

Task 7

Determine the number of different ways to partition a set of n elements into k clusters.

Task 8

What are the differences between EM and k-means?

References

1. Gan Guojun, Chaoqun Ma, Jianhong Wu. Data clustering: theory, algorithms, and applications. ASA-Siam Series on Statistics and Applied Probability; 2007.
2. Pfeiffer U. Dirk, Robinson P. Timothy, Stevenson Mark et al. Spatial Analysis in Epidemiology. NY: Oxford University Press; 2008.
3. Waller LA., Jacquez GM. Disease models implicit in statistical tests of disease clustering. Epidemiology. 1995;6:584-90.
4. Ritter G., Gallegos M. T. Outliers in statistical pattern recognition and an application to automatic chromosome classification. Pattern Recognition Letters. 1997;18:525–39.
5. Rezaeian M, Dunn G, Leger St, Appleby L. Geographical epidemiology, special analysis and geographical information systems: a multidisciplinary glossary. J. Epidemiol. Community Health. 2007;61:98-102.
6. Sjurdur FO, Martuzzi M, Elliot P. Cluster analysis and disease mapping- why, when, and how? A step by step guide. Br Med J. 1996;313:863-6.
7. Knox EG. Detection of clusters. In: Elliot P. ed. Methodology of enquiries into disease clustering. London: Small Area Health Statistic Unit. 1989;17-20.
8. Elliott P, Martuzzi M, Shaddick G. Spatial statistical methods in environmental epidemiology. Statistical Methods in Medical Research. 1995;4:139-61.
9. Jyun-de Wu, Milton DK, Hammond SK, Spear RC. Hierarchical Cluster Analysis Applied to Workers Exposures in Fiberglass Insulation Manufacturing. Ann. Of Occupational Hygiene. 1999;43:43-55.
10. Kaufman L, Rousseeuw PJ. Finding groups in data: An introduction to cluster analysis. NY:J. Wiley &Sons; 1990.
11. Kiers H.A.L et al. (Eds.): Data Analysis, Classification, and Related Methods. Berlin: Springer. 2000;53-8.

12. Bezdek JC, Keller J, Krisnapuram R, Pal NR. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. Dordrecht: Kluwer; 1999.
13. Jain AK, Dubes RC. Algorithms for Clustering Data. Englewood Cliffs, NJ: Prentice-Hall; 1988.
14. Aldenderfer M, Blashfield R. Cluster Analysis. Beverly Hills, CA: Sage Publications; 1984.
15. Witten IH, Frank E. Data Mining: Practical Machine learning Tools and Techniques. 2nd ed. Elsevier; 2005.
16. Little R, Rubin D. Statistical Analysis with Missing Data. NY: J. Wiley & Sons; 1987.
17. Hartigan JA. Clustering Algorithms. NY: J. Wiley & Sons; 1975.
18. Everitt B, Landau S, Leese M. Cluster analysis. 5th ed. UK: J. Wiley & Sons; 2011.
19. Kendall M. Multivariate Analysis. 2nd Edition. NY: Macmillan; 1980.
20. Mardia K, Kent J, Bibby J. Multivariate Analysis. Academic Press; 1979.
21. Massart D et al. Non-hierarchical clustering with MASLOC, J Pattern Recognat Society. 1983;16:507-16.
22. Banks D, House L, McMorris FR, Arabie P, Gaul W. (eds). Classification, Clustering and Data Mining Applications. Berlin: Springer-Verlag; 2004.
23. Cherkassky V, Mulier F. Learning from data: concepts, theory and methods. NY: J. Wiley & Sons; 1998.
24. Michie D, Spiegelhalter D, Taylor C. (eds). Machine Learning, Neural and Statistical Classifications. NY: Ellis Horwood; 1994.
25. Milligan GW, Cooper MC. A study of standardization of variables in cluster analysis. J Classificat. 1988;5:181-204.
26. Tango Toshiro. Statistical Methods for Disease Clustering. Springer; 2010.
27. Gordon AD. Boca Raton FL. Classification. London: Chapman and Hall/CRC; 1999.
28. Hand D. Discrimination and Classification. Chichester: J. Wiley & Sons; 1981.
29. Forgy E. Cluster analysis of multivariate data: Efficiency vs. interpretability of classifications. Biometrics. 1965;2:768-9.
30. McLachlan GJ. Discriminant Analysis and Statistical Pattern Recognition. NY: J. Wiley & Sons; 1992.
31. Hartigan JA., Wong MA. A k-means clustering algorithm. Applied Statistics. 1979;28:100-8.
32. Coleman DA., Woodruff DL. Cluster analysis for large datasets: An effective algorithm for maximizing the mixture likelihood. J. Comput. Graph. Statist.2000;9:672-88.
33. Barnett V, Lewis T. Outliers in Statistical Data. 3rd ed. Chichester: J. Wiley & Sons; 1994.
34. Wild CJ, Seber GAF. Chance Encounters: A first course in data analysis and inference. NY: J. Wiley and Sons; 2000.
35. Jordan M, Jacobs R. Hierarchical mixtures of experts and the EM algorithm. Neural Comput. 1994;6:181-214.
36. Neal R., Hinton G. A view of the EM algorithm that justifies incremental, sparse, and other variants. In: JKordan M, editor. Learning in Graphical Models. Dordrecht: Kluwer Acad. Publ. 1998;355-68.
37. Zhang P. Model selection via multifold cross-validation. Ann. Statist. 1993;21:299-311.

Recommended readings

1. Everitt B, Landau S, Leese M. Cluster analysis. 5th ed. UK: J. Wiley & Sons; 2011.
2. Jyun-de Wu, Milton DK, Hammond SK, Spear RC. Hierarchical Cluster Analysis Applied to Workers Exposures in Fiberglass Insulation Manufacturing. Ann. Of Occupational Hygiene. 1999;43:43-55.
3. Gan Guojun, Chaoqun Ma, Jianhong Wu. Data clustering: theory, algorithms, and applications. ASA-Siam Series on Statistics and Applied Probability; 2007.
4. Rezaeian M, Dunn G, Leger St, Appleby L. Geographical epidemiology, special analysis and geographical information systems: a multidisciplinary glossary. J. Epidemiol. Community Health. 2007;61:98-102.
5. Sjurdur FO, Martuzzi M, Elliot P. Cluster analysis and disease mapping- why, when, and how? A step by step guide. Br Med J. 1996;313:863-6.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Introduction to intervention studies**
Module: 2.16	ECTS (suggested): 0.1
Author(s), degrees, institution(s)	Tatjana Pekemezovic MD, PhD, Professor, Institute of Epidemiology, Faculty of Medicine, University of Belgrade, Serbia Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Tatjana Gazibara, MD, Teaching Assistant, Institute of Epidemiology, Faculty of Medicine, University of Belgrade, Serbia
Address for correspondence	Tatjana Pekmezovic Institute of Epidemiology, Faculty of Medicine, University of Belgrade, Visegradska 26A, Belgrade, Serbia E-mail: pekmezovic@sezampro.rs
Keywords	Clinical trial, community trial, experimental study, field trial Intervention study, randomization.
Learning objectives	At the end of this topic student will be able to: • describe and explain basic concepts of intervention studies; • understand and explain different types of intervention studies; • participate in study design, derivation and data analysis; • read epidemiological literature that use and refer to the concepts outlined above.
Abstract	First, a distinction between observational and experimental studies is made. Afterwards, current ideas and trends in experimental epidemiology are explored, as well as basic characteristics of different types of intervention studies. During this topic key interrelated components in intervention studies, such as planning, organization, selection of study participants, calculation of sample size, follow-up, and detection of the effects, are introduced. The potential biases in study design and measurement of outcomes make an important part of this topic, too. In addition, an ethical consideration of intervention studies is discussed.
Teaching methods	The teaching method recommended: • the introductory lecture related to topics mentioned above; • exercises which include distribution of selected papers (different types of intervention studies) to each student and their critical appraisal (discussion in small groups); • development of a written protocol (project proposal) of an intervention study according to problem assigned in small group (up to 5 students); • presentation of project proposals and overall discussion related to advantages and disadvantages of each design.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire and short written essay according to assigned study problems.

INTRODUCTION TO INTERVENTION STUDIES

Tatjana Pekmezovic, Lijana Zaletel-Kragelj, Tatjana Gazibara

Theoretical background

Introduction

Previous modules in this book were dealing with different types of observational studies while in this module the experimental designs are briefly introduced.

Basic definitions

Prior discussing the characteristics of this group of epidemiological studies it could be worthy to give some definitions of terms frequently used in relation to experimental studies.

1. An experiment
   Among definitions of the term "experiment" we can find the following:
   - according to Rothman et al. (1), an experiment is a set of observations, conducted under controlled circumstances, in which the scientists manipulates the conditions to examine/verify what effect such a manipulation has on the observations. In epidemiology, the term "experiment" usually means that the investigator manipulates the exposure assigned to participants in the study. Usually, an experiment refers to any trial or test.
2. An experimental study
   Among definitions of the term "experimental study" we can find the following:
   - according to A dictionary of epidemiology (2), experimental study is a study in which conditions are under the direct control of the investigator. In epidemiology it is a study in which a population is selected for a planned trial of a regimen whose effects are measured by comparing the outcome of the regimen in the experimental group with the outcome of another regimen in a control group. As examples of experimental studies, a randomized controlled trial and a community trial are mentioned;
   - according to the TheFreeDictionary's Medical dictionary (3), an experimental study is a study in which all of the risk factors are under the direct control of the investigator.
3. An intervention
   Among definitions of this term we can find the following:
   - according to Beaglehole et al. (4), an intervention or experimentation is an act that involves attempting to change a variable (an outcome) in one or more groups of people. The effects of an intervention are measured by comparing the outcome in the experimental group with the outcome in a control group;
   - according to Medicine.Net Online Dictionary (5), an intervention is the act of intervening, interfering or interceding with the intent of modifying the outcome. In medicine, an intervention is usually undertaken to help treat or cure a condition. It comes from the Latin term "intervenire", meaning "to come between";
   - according to the TheFreeDictionary's Medical dictionary (3), an intervention is a) the act or fact of interfering so as to modify, and b) any measure whose purpose is to improve health or alter the course of disease.
4. An intervention study
   Among definitions of this term we can find the following:
   - according to A dictionary of epidemiology (2), intervention study is an investigation involving intentional change in some aspects of the status of the subjects, e.g. introduction of a preventive or therapeutic regimen, or designed to test a hypothesized relationship; usually an experiment such as a randomized controlled trial;
   - according to the TheFreeDictionary's Medical dictionary (3), an intervention study is a testing of a hypothesized epidemiological cause-effect relationship by intervening in a population and modifying a supposed causal factor and measuring the effect of the change.

   In fact, the term "intervention study" is an alternative for the term "an experimental study" (4).
5. To control
   Among definitions of this term we can find the following:
   - according to A dictionary of epidemiology (2), the term "to control" among others means "to regulate" (in terms of experimental studies this means that a researcher gains the mastery over the situation in opposite to observational studies where a researcher only observe the situation/process and do not intervene).
6. A trial
   Among definitions of this term we can find the following:

- according to the TheFreeDictionary's Medical dictionary (3), the term "trial" is a synonym for the term "experiment"; it refers usually to the trying out of a substance or a material in order to determine its effect;
- the same states Aschengrau and Seage (6).

Description of intervention studies

General characteristics of intervention studies

Intervention (experimental) studies, also known as intervention trials (7), are the epidemiological studies that are most similar to laboratory experiments. They are characterized by the property that in these studies an investigator directly controls on experimental circumstances, and more precisely, intercedes with new therapeutic agent, vaccine, or preventive procedure (7,8). As we mentioned previously, the effects of interventions are measured by comparing the outcome in both experimental and control groups.

The major distinction between observational and intervention studies underlies in the fact that in observational studies the investigator accepts the conditions as they are and makes observations with aim to answer questions related to studied problem, while in intervention studies he/she gains the mastery over the situation.

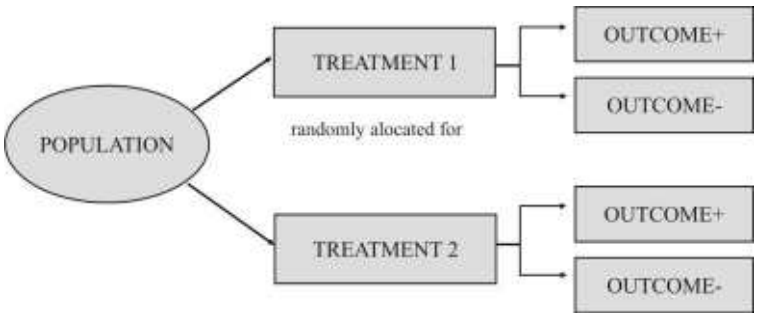
If treatments are allocated randomly in a sufficiently large sample, intervention studies have the potential to provide a degree of certainty about the validity of results that is absolutely not possible with any observational study (9). In other words, intervention studies provide the strongest evidence with which to test hypothesis. In addition, it is considered to be the ideal design for evaluating the effectiveness and effects of new treatment or intervention. Besides that, intervention studies can be used for many purposes: evaluating new drugs or other treatments, testing of new health and medical care technology, assessment of new programs for screening, or finding new ways of organization and delivery health services (10). However, due to mainly ethical and practical reasons, this approach is relatively a rare study design in epidemiology and public health (11). Other limitations include difficulties in generalization of results, limited feasibility, response and attrition problems, extreme expenses because of large number of participants and engaged investigators, etc.

Randomization - the essential characteristic of intervention studies

The essential characteristic of intervention studies is the randomization, although an experiment can also be non-randomized what will be discussed later in this section.

Randomization (or random allocation) of individuals (or groups or populations) is a process of allocation (assignment) of participants in the study (individuals or communities) to the experimental and the control group (Figure 1) by chance (2). After the random allocation the intervention procedure is applied to the experimental, but not to the control group. After the follow-up period is completed the effect is assessed according to previously defined outcome (11).

Figure 1. General design of intervention (experimental) studies. Adapted from Dos Santos Silva (7)



Within the limits of chance variation, randomization is intended to make the control and the experimental groups similar at the start of the trial (2). In other words, it eliminates selection bias on the part of the participants and investigators. Because of this property it is one of the methods of controlling for the potential confounding factors (7).

Various methods can be used to randomize the study subjects to different study groups such as:

- simple randomization - the most elementary method of randomization, equivalent of tossing a coin (7). Selection of the subjects occurs randomly. Randomization list could be obtained by using a table of random numbers, or it could be computer-generated by using random number generator;
- stratified randomization, including the match-pair design (frequently used in a community trials) - used when the results of the trial are likely to vary between categories of certain characteristic (i.e. between sexes/genders or between different age groups). First strata are formed and randomization occurs separately for the subjects in each stratum (7);
- cluster randomization - this method is characterized by the property that randomization occurs at the group (cluster) level. All individuals within a given cluster are assigned to the same study group (so-called arm).

Cluster sampling is typically used when the researcher cannot get a complete list of the members of a population they wish to study but can get a complete list of groups or "clusters" of the population. It is also used when a random sample would produce a list of subjects so widely scattered that surveying them would prove to be far too expensive, for example, people who live in different districts (12).

Cluster randomized trials are less efficient statistically than individually randomized trials because the responses of individuals in a cluster tend to be more similar (intracluster) than those individuals in different clusters (intercluster). The sample size required is accordingly larger and the analysis techniques have to be adjusted by the level of association among members of the cluster (intracluster correlation coefficient) (13).

Despite the randomization is the most optimal design for evaluating effectiveness, in practice (14), conducting a randomized controlled trial is not always feasible. Some alternatives to randomization include historical controls and non-randomized controls.

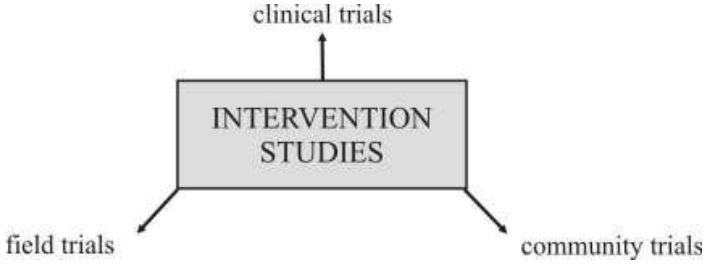
- Historical controls can be used in case that we have therapy today that we believe will be quite effective, and would like to test it in a group of patients. Therefore, for comparison, we will go back to the records of patients with the same disease who were treated before new therapy became available. This type of design seems inherently simple and attractive, but we can not be sure that the differences between such groups are due to the therapy, because many things other than therapy change over time.
- One of alternative approaches is also to use controls that are not selected in a randomized manner. In such a case we are talking of non-randomized trials**.**
  In the case of clinical trials**,** non-randomized trial is a clinical trial in which the participants are not assigned by chance to different treatment groups. Participants may choose which group they want to be in, or they may be assigned to the groups by the researchers (15).
  The essential characteristic of non-randomized trials is that the participants are not assigned by chance to different treatment groups. Participants may choose which group they want to be in, or they may be assigned to the groups by the researchers.
  Despite the fact that non-randomized trials do not yield the same sort of information tas the randomised trials, they have its own importance - the purpose is exploratory or hypotheses-generating. On the basis of their results something must be subsequently proved by randomised trials (14).
  Non-randomized trials are often used to evaluate the effectiveness of surgical treatments (16). In this type of intervention study, the three aspects of design are the following:
  - the relationship in time between groups being compared,
  - whether or not participants were treated before the study was conceived, and
  - the basis for allocating a treatment to a participant.

These characteristics are important because they are believed to influence the risk of bias. It is well known that in randomized controlled trials, treatment is allocated by chance, while in non-randomized studies, the decision to give a particular treatment to a particular participant is made by researchers. Both methods are likely to give rise to imbalances in prognostic factors between the groups being compared. Imbalances can be reduced by matching, or controlled by statistical methods, but their confounding effects can never be completely removed (7). Imbalances can bias estimates of treatment effects and almost always increase their uncertainty (16).

Types of intervention studies

Intervention studies (epidemiologic experiments) include (1,4): clinical trials, field trials, and community trials (community intervention studies) (Figure 2).

Figure 2. Three types of intervention studies



This classification in fact comprises two classifications (6):

- clinical trial versus field trial, and
- individual trial versus community trial.

Clinical trials are trials performed on individual patients as subjects of investigation (mostly in hospitals/clinics); field trials are trials performed on individual (healthy) community members as subjects of investigation; whereas, community trials are performed on whole communities as subjects of investigation.

This is only one of possible classifications of intervention studies. Here are some others (6):

- therapeutic trial versus prophylactic trial,
- randomized trial versus non-randomized trial,
- simple trial versus factorial trial, etc.

Clinical trials

Definition

Clinical trials are:

- according to the TheFreeDictionary's Medical dictionary (3), experiments performed on human beings in order to evaluate the comparative efficacy of two or more therapies;
- according to the U.S National Institutes of Health (17), clinical trials are studies to answer specific questions about vaccines or new therapies or new ways of using known treatments;
- according to Rothman et al. (1), experiments with patients as subjects, with a goal to evaluate a potential cure for disease, or to find a preventive of disease sequelae such as death, disability, or a decline in the quality of life.

Randomized controlled (clinical) trials

Clinical trials are mostly randomized. In this case are referred to as randomized controlled (clinical) trials (RCTs):

- according to Lilienfeld and Stolly (8), randomized controlled (clinical) trials are epidemiologic experiments mainly conducted with the aim to study the efficacy of a drug or a medical procedure (e.g. surgical intervention) in the treatment of a disease. However, as Lilienfeld and Stolly state, they can also be used to evaluate a preventive agents (such as vaccines) or public health procedures (for example, screening) (8). While the different interventions are in question, the general methods and principles, for the most part, remain the same,
- according to Last et al. (2), RCTs are epidemiologic experiments in which subjects in a population are randomly allocated into a study and a control group to receive or not an experimental preventive or therapeutic procedure (or intervention).

To ensure that the groups being compared in RCTs are equivalent, patients are assigned to them randomly. RCTs are considered as the most scientifically rigorous method of hypothesis testing available in epidemiology (2).

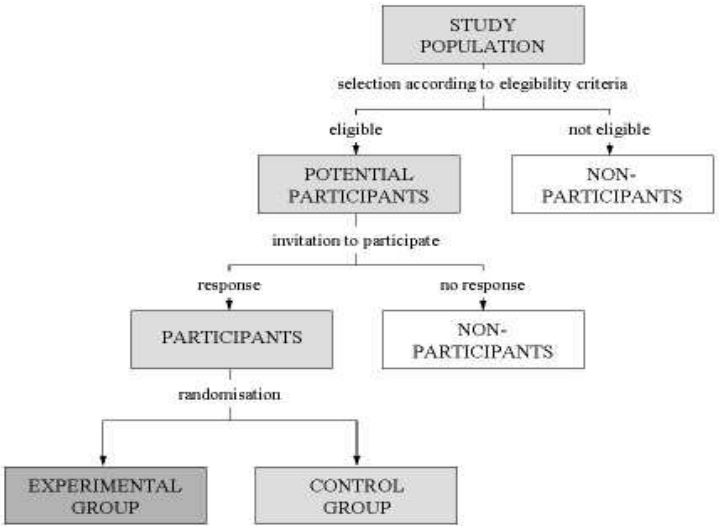
Clinical trials designs

Clinical trials can be conducted according to different designs (6,7,14,18,19): parallel groups, cross-over, or factorial design.

1. Parallel group design (6,18,19)

   The most common design is the parallel group design in which patients are randomised to one of two or more so-called arms of a clinical trial. One arm is being allocated to an experimental group and the other to a control group. There exist several types of comparison of experimental group (an intervention under observation) to a control group (20). For example, it could be compared to placebo, another intervention, same intervention at a higher dose (or longer duration), or no intervention. This design of clinical trials is schematically presented in Figure 3 (4).

Figure 3. General design of randomized controlled (clinical) trials. Adapted from Beaglehole et al. (4)



2. Cross-over design (6,7,14,18,19)
   The cross-over design is one of the within-patient comparison designs. In this design, each patient is randomised to a sequence of two or more treatments, and thus acts as his/her own control for treatment comparisons. This design is attractive primarily because it reduces the number of patients required to achieve a specific power, sometimes to a marked extent. In the simplest two-by-two cross-over design, each patient receives each of the two treatments in randomised order in two successive treatment periods, often separated by a washout period.
3. Factorial design (6,7,14,18,19)
   In a factorial design, two or more treatments are evaluated simultaneously in the same patient population through the use of varying combinations of the treatments. In other words, each group of patients gets two or more treatments.
   The simplest example is the two-by-two factorial design in which patients are randomly allocated to one of the four possible combinations of two treatments. If these two treatments are labelled as treatment A and treatment B, the combinations are: A alone; B alone; both A and B; neither A nor B.
   The usual intention of using factorial design is to make efficient use of clinical trial patients by evaluating the efficacy of the two treatments with the same number of patients as would be required to evaluate the efficacy of either one alone. In other words, factorial design allows answering to more questions in a single trial for minor increase in costs.

There exist also some other designs, namely match pairs design, sequential design and other within-patient comparison designs like Latin and Greco-Latin square designs (14).

Detailed discussion on strengths and limitations of various designs are out of the scope of this module.

Clinical trials at different phases of experimental clinical research

Experimental clinical research usually progresses in an orderly series of steps, called phases. The trials at each phase have a different purpose and help scientists answer different questions (14,21,22):

1. Phase I trials
   By these trials, researchers test an experimental drug or treatment for the first time. Main goals are to detect potentially harmful adverse effects of observed treatment, and to determine the metabolic and pharmacological actions, and safe dosage range. The observed group of participants is small (20-80). It is preferably to recruit healthy volunteers, because an unexpected and potentially dangerous reaction can occur, that is easily manageable in healthy participants. Duration of these studies is up to 1 month.
2. Phase II trials
   During this phase the potential treatment's therapeutic usefulness/effectiveness is evaluated. It is aimed also at determining the short-term side effects, and to identify common risks for a specific population and disease. The experimental study drug or treatment is administered for a limited period (several months) to a smaller number of patients with target disease (100-300) that must be as homogenous as possible. This trial is usually offered to patients who have not improved with other available treatments.
3. Phase III trials
   In phase III trials, the experimental study drug or treatment is extended to larger and less homogenous group of patients with the target disease (1,000-3,000) to confirm its effectiveness, monitor side effects, compare it to commonly used treatments, and collect information that will allow the experimental drug or treatment to be used safely. Patients involved in this phase are patients with the same type of cancer who otherwise would receive best current treatment. The duration of administration increases (several years).
4. Phase IV trials
   In phase IV trials, post-marketing studies provide additional information including the drug's risks (long-term side effects), benefits (additional uses of the agent), and optimal use. The number of participants increases (thousands of individuals with target disease as well as new population groups), as well as the observation period (on-going process).

Types of clinical trials

There exist several types of clinical trials. According to classical classification by Lilienfeld and Stolly (8), there exist three types of clinical trials, being therapeutic, intervention and preventive clinical trials:

- **Therapeutic trials**

Therapeutic trials are carried out with the aim to cure diseases, prevent recurrences and complications or increase survival. Subjects with a disease are involved in the study. An example of such a study is a study on a Zidovudine (AZT) treatment for acquired immunodeficiency syndrome (AIDS) (23) (Example 1).

Example 1. AZT is a potent inhibitor of the replication of the human immunodeficiency virus type 1 (HIV), and it has been shown to improve survival in advanced HIV disease.

The objective of this study was to evaluate the efficacy and safety of AZT, early in the treatment of HIV infection.

A double-blind, randomized, placebo-controlled trial was performed, with subject stratification by pre-treatment CD4 T lymphocyte counts. It was a multicentre trial at AIDS Clinical Trial units in the USA. Included in the study were seven hundred eleven subjects with mildly symptomatic HIV infection. Three hundred fifty-one subjects were assigned to placebo and 360 to AZT, 200 mg orally every 4 hours. The median duration of follow-up was 11 months.

Fifty-one subjects developed the AIDS, advanced AIDS-related complex, or death as a first critical event. For the stratum of subjects with more than 200 but less than 500 CD4 T lymphocytes/mm3 before treatment, 34 events occurred in placebo recipients and 12 in AZT recipients (p=0.0002). For the stratum of subjects with 500-799 CD4 T lymphocytes/mm3 before treatment, 2 events occurred in placebo recipients and 3 in AZT recipients. Significant differences between the treatment groups in CD4 T-lymphocyte counts occurred in subjects with more than 200 but less than 500 CD4 T lymphocytes/mm3 after 4 weeks of therapy (p=0.002). Differences persisted through week 52. Less prominent changes occurred in subjects with 500 or more CD4 T lymphocytes/mm3. Serum levels of HIV antigen decreased significantly in AZT recipients. Serious anaemia and neutropenia occurred in 5% and 4% of AZT recipients, respectively, and in 0% and 1% of placebo recipients, respectively.

In conclusion, AZT delayed progression of HIV disease and produced little toxicity in subjects with mildly symptomatic HIV disease and less than 500 CD4 T lymphocytes/mm3.

- **Intervention trials**

In intervention trials investigator intercedes before a disease has developed in subjects at high risk of getting a disease. An example of such a study is s study on an AZT treatment of HIV-positive individuals without AIDS symptoms (24) (Example 2).

Example 2. Since AZT has been shown to improve survival in advanced the human immunodeficiency virus (HIV) disease, the aim of present study was to estimate efficacy and safety of this drug in persons with asymptomatic HIV infection.

A randomized, double-blind trial was conducted in adults with asymptomatic HIV infection who had CD4+ cell counts of fewer than 500 per cubic millimetre on entry into the study. The subjects (92 percent male) were randomly assigned to one of three treatment groups: placebo (428 subjects); AZT, 500 mg per day (453); or AZT, 1500 mg per day (457). A mean follow-up was 55 weeks.

After a follow-up period, 33 of the subjects assigned to placebo had AIDS, as compared with 11 of those assigned to receive 500 mg of AZT (p=0.002) and 14 of those assigned to receive 1500 mg of AZT (p=0.05). In the three treatment groups, the rates of progression (per 100 person-years) to either AIDS or advanced AIDS-related complex were 7.6, 3.6, and 4.3, respectively. As compared with those assigned to placebo, the subjects in the AZT groups had significant increases in the number of CD4+ cells and significant declines in p24 antigen levels. In the 1500-mg AZT group, severe haematologic toxicity (anaemia or neutropenia) was more frequent than in the other groups (p less than 0.0001). In the 500-mg AZT group, nausea was the only toxicity that was significantly more frequent (in 3.3 percent) than in the placebo group (p=0.001).

The authors concluded that AZT is safe and effective in persons with asymptomatic HIV infection and fewer than 500 CD4+ cells per cubic millimetre.

- **Preventive (prophylactic) trials.**

Preventive trials are conducted with aim to estimate the efficacy of a preventive agent or procedure among subjects free of disease. An example of such a study is a study on an education in use of condoms in prevention of HIV transmission and infection (25) (Example 3).

Example 3. In some parts of Africa, prostitutes and their clients represent the groups at greatest risk of human immunodeficiency virus (HIV) infection and the major disseminators of the virus. Condom use was assessed after a programme of education about the AIDS (acquired immunodeficiency syndrome) and a condom distribution programme in a well-characterised prostitute population in Nairobi. Women received their education at group meetings (barazas) and at individual counselling sessions during which they were given the results of serological tests for HIV (group 1) or at barazas only (group 2), or through very little of either (group 3). During the counselling sessions free condoms were distributed. Before either of the programmes started, 10%, 9%, and 7% of groups 1, 2, and 3 women, respectively, reported occasional use of condoms. During the first year of study, 80%, 70%, and 58% of groups 1, 2, and 3 women, respectively, reported at least some condom use. The mean frequency of condom use was 38.7%, 34.6%, and 25.6% of sexual encounters in groups 1, 2, and 3 women. 20 of 28 women who were non-condom-users seroconverted compared with 23 of 50 women who reported some use of condoms.

Beside the presented classification, also other more recent classifications exist. For example, according to the U.S. National Institutes of Health (21) there are five types of clinical trials, being treatment, prevention, diagnostic, screening, and quality of life clinical trials:

- treatment trials test new drugs or new combinations of drugs, or new therapeutic approaches (i.e. new approaches to surgery or radiation therapy;
- prevention trials test new approaches, such as medicines, vitamins, minerals, or other supplements that are believed that may lower the risk of a certain types of diseases (i.e. cancer), or lifestyle changes. In other words

they look for better ways to prevent disease in people who have never had the disease or to prevent a disease from returning;

- diagnostic trials are conducted to find better tests or procedures for diagnosing a particular disease or condition,
- screening trials test the best way to detect certain diseases or health conditions (i.e. cancer, especially in its early stages);
- quality of life trials (or supportive care trials) explore ways to improve comfort and the quality of life for individuals with a chronic illness.

Field trials

Field trials, in contrast to clinical trials, as a rule, deal with subjects who are free of disease but presumed to be at risk, and involve evaluation of agent or procedure with the aim to reduce the risk of developing disease in general population.

Data collection takes place "in the field", usually among non-institutionalized people in the general population (4).

This design usually requires a larger number of subjects and longer follow-up period than clinical trials, since their purpose is to prevent the occurrence of diseases that typically occur with relatively low frequency (1,4,7). Additionally, since subjects are not under active health care in health care settings (e.g. under treatment in community health centre or even hospitalized in hospitals), they do not come to a central location for treatment. Consecutively, field trials often require visiting subject at home or on the work-place (school-place), or establishing study centres. All these characteristics mean that field trials are huge projects involving a lot of human and financial resources (1).

A random allocation of individuals to an experimental and a control group is again an ideal design. However, in practice there are a lot of difficulties for its implementation (1). Consecutively, other designs are frequently applied (e.g. cluster randomization). The problem is that these modifications can affect the informativeness and interpretation of experimental findings (1). Detailed description of these limitations is out of the scope of this module.

An example of such a study is a study on the effect of breast cancer screening on mortality from breast cancer (26) (Example 4).

Community trials

Trying to find what community trials are we can find the following definitions:

- according Rothman et al. (1), community intervention trials are an extension of the field trials that involves intervention on a community level,
- according to Dos Santos Silva (7), community trials are a special form of field trials in which whole communities are the unit of allocation. Community trials thus involve population as a whole, i.e. the group as a whole studied collectively (7),
- according to Last et al. (2), a community trial is an experiment in which the unit of allocation to receive a preventive or therapeutic regimen is an entire community or political subdivision.

Example 4. With the aim to evaluate the effect of breast cancer screening on mortality from this disease, women, aged 40-60 years who were members of the Health Insurance Plan of town X, were randomly divided into two groups: intervention group (four mammography examinations were offered at annual intervals) and control group (receiving usual medical care). Each group comprised 31,000 women. The groups were very similar in terms of demographic and other characteristics of interest (26).

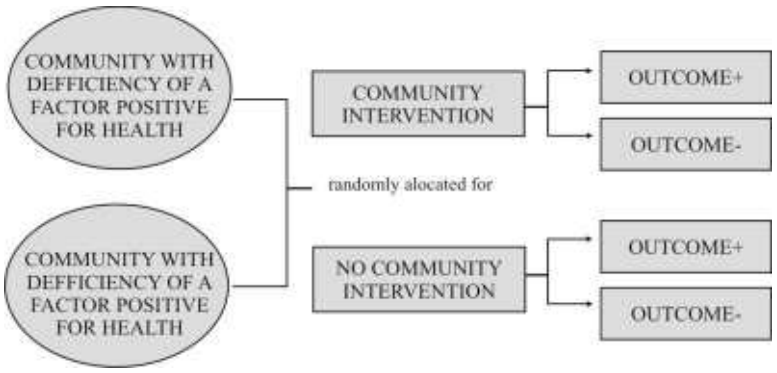
A critical factor in determining sample size was the interest in detecting at least a 20% reduction in mortality that might be attributed to screening. High levels of comparability between the study and control groups have been demonstrated for a wide range of demographic and other characteristics and for general mortality other than breast cancer.

Screening consisted of a clinical examination, usually by a surgeon; mammography, in which two views were taken of each breast (cephalo-caudal and lateral); and an interview to obtain relevant demographic information and a health history. Independence between the two examining modalities was strictly maintained so that it could be determined which was responsible for the chain of events that led to biopsy. Control women continued to receive their usual medical care.

By the end of 10 years after entry, the study group's mortality due to breast cancer was about 30% below the control group. However, there was no longer difference in mortality from causes other than breast cancer between intervention and control groups.

General design of community intervention studies is similar to the general design (parallel) in clinical trials (Figure 4). Again, a random allocation of study units (groups, communities) to an experimental and a control group is again an ideal design (in this case, cluster randomization).

Figure 4. General design of community intervention studies



An example of such a study is the Newburgh-Kingston caries fluorine study (27) (Example 5). Another well-known example is a trial of heart disease prevention in North Karelia, Finland (2,28).

Example 5. The aim of Newburgh-Kingston dental caries study was to estimate whether the increased fluoride concentration in drinking water might reduce prevalence of decayed, missing or filled permanent teeth. One entry community (Newburgh) was allocated randomly to receive fluoride added to the water supply, while the other (Kingston) continued receiving water without supplementation (27).

The caries fluorine hypothesis which states that fluorine has a prophylactic effect on dental caries is supported by extensive epidemiological studies in the USA and in other parts of the world.

The aim of Newburgh-Kingston dental caries study was to estimate whether the increased fluoride concentration in drinking water might reduce prevalence of decayed, missing or filled permanent teeth.

In 1944, the study plan was made a reality when the cities of Newburgh and Kingston in New York State agreed to participate in such a program as study and control areas, respectively. This study was started in June, 1944, when basic dental examinations were begun. One entry community (Newburgh) was allocated randomly to receive fluoride added to the water supply, while the other (Kingston) continued receiving water without supplementation. Both cities are situated on the Hudson River about 30 miles apart. Each has a population of approximately 30,000. The climate of both cities is also similar, and their water supplies at the outset of this study were comparable and have remained so, except for the addition of sodium fluoride to Newburgh's supply.

On May 2, 1945, sodium fluoride was added to Newburgh's water supply to bring its fluorine content up to 1.0-1.2 p.p.m., while Kingston's water supply remained fluorine-free.

All dental examinations in Newburgh and the first series in Kingston were made with mouth mirror and sharp explorer by the same examiner. The subsequent examinations in Kingston, using the same technique, were made by two dental hygienists trained in the method of examination and the charting of defects. In both areas the examiners called off the defects which were recorded by a staff clerk on a dental record card designed specifically for this study.

It was expected that the study would take 10-12 years to determine adequately the efficacy and safety of this caries prophylactic measure.

The proportion of erupted permanent teeth with evidence of caries experience (decayed, missing, or filled) decreased in each successive examination period in Newburgh, from 21 per 100 before water fluoridation to 14.8 per 100 at the time of the last survey. This rate in Kingston remained approximately 21 for the examination period. The difference between Newburgh and Kingston at the last examination suggested a 30% improvement in Newburgh.

Course of intervention studies

Study protocol

Before the start of the study, protocol in written form must be developed. It contains rationale and specific objectives, precise description of methods for selecting and allocating study groups, number of participants, randomization schemes, criteria of including and excluding participants, type and duration of intervention, major and minor outcomes, as well as methods for their monitoring and registration. In addition, methods for data collection and analysis, and procedures for obtaining the informed consent of subjects must be described. The protocol should be accompanied by all forms which will be completed during the study (usually, as appendix).

Selection of participants

The next step is selection of study population. Participants must be similar in regard to many characteristics that could significantly influence the outcome, such as sex, age, degree of previous exposure, stage of disease or its absence, etc. Investigator selects target population, which comprises individuals or groups with a set of characteristics related to the problem investigated. When the target (or reference) population is defined, one needs to select the actual population in which the study will be conducted. It is the experimental population. Inclusion and exclusion criteria must be clearly defined before the study begins. It is very important, especially as

a method for elimination of all unsuitable subjects (mainly those with characteristics which can interfere with the outcome). Eligible subjects must be invited to participate in the study, after being fully informed about the purposes, procedures, possible risks and benefits of the study. After exclusion of refusals, the study population is defined. In community trials selected communities should be stable, with little migration and have self-contained medical care system (8).

Sample characteristics

The sample size, defined as a number of individuals, necessary to detect the effect of intervention, is an essential part of experiment preparation. Inadequate samples can cause lack of improvement, which exists and is confirmed with large sample (8). The sample size is computed applying various statistical procedures. The following steps in sample size calculation are generally accepted:

- detection of difference in response rates
- estimation of the response rate in one of the groups
- detection of level of statistical significance (alpha)
- detection of the value of the power desired (1-beta), and
- detection whether the test should be one-sided or two-sided.

In circumstances when a rare type of exposure or outcome is in question, a sufficient number of participants may pose a big problem. In such situation, multicentric trial or meta-analysis can be useful. In multicentric trials, many hospitals (or groups) in the community, country, or throughout the world are included in the study using the same study protocol. Meta-analysis is a method in which data from similar studies are pooled in a statistically rigorous manner (9).

Randomization

The next step in the design of intervention study is allocation of participants into the test and comparison groups with the aim to ensure that those treated and those untreated are exactly similar in almost all aspects before intervention. It is the best method if assignment to study groups is done at random.

Compliance

The important issue of experimental design is monitoring compliance and side effects. Noncompliance can decrease the statistical power of study to detect exact effect of intervention, although a certain degree of noncompliance is acceptable, especially in estimation of effectiveness of intervention in real-life conditions.

Patients may be randomized, but following randomization they may not comply with the assigned treatment. Noncompliance may be overt or covert. In the first case, people may overtly articulate their refusal to comply or may stop participating in the study. These non-compliers are dropouts. On the other side, people may just stop taking the agent assigned without admitting this to the investigator or the study staff. Another problem in clinical trials is drop-ins. In this case, patients in one group may inadvertently take the agent assigned to the other group.

Outcome assessment

The outcome of interest must be clearly defined before the study commencement. In clinical trials, effects should be evaluated in each patient. In community trials, outcomes would be expressed as a reduction of incidence of disease or cost to health services (11). If investigator knows whether participants were in treatment or control group it can result with a biased assessment of effect. To eliminate this problem, three procedures have been developed: single-, double- or triple-masking. In a single-masking study, participants are not given any indication to whether they belong to treatment or control group. The aim is to prevent participants from introducing bias into observations; it can be achieved by use of a placebo. In a double-masking study, neither participants nor investigator have knowledge of the participant group allocation. In a triple-masking study, participants, investigator and reviewer of data are all masked with regard to the group individuals belong to (8).

Measurements of outcome include both improvement (the desired effect) and any side effects that may appear. Therefore, there is a need for explicitly stated criteria for all outcomes to be measured in a study.

Follow-up

Procedures during the follow-up period are the same for all study participants. In this part of intervention study important issues include equal and rigorous follow-up in both groups, simple but sufficient methods for detecting of all relevant events, and high quality cooperation (in this way, loss from the study population should be minimized) (11).

Data analysis

The data analysis is performed with the aim to assess the efficacy of intervention. For example, in vaccine trials, the efficacy is the proportion (or percentage) of the expected incidence of disease which is prevented by intervention. In case when observed benefits are high or possible injury effects are serious, results must be analyzed sequentially. This means to continue data analysis and stop the study when a significant benefit or adverse effect has been demonstrated (11).

Ethical considerations

In intervention studies, since investigator deliberately intervenes, ethical considerations are more important in comparison to other types of epidemiological studies. Before the study is carried out, many questions should be considered. Hill mentioned some of them such as: whether the proposed treatment is safe or possibly harmful for study participants, whether it is ethical to use placebo treatments, etc. (29). As mentioned above, each participant in the study must be fully informed about the purposes and potential adverse effects of intervention. If subjects provided with this information decide to participate, their informed consent must be obtained. Personal privacy and confidentiality must be respected at all time. Nowadays, almost all research and health institutions have Ethical Committees, formed with the aim to control and monitor ethical aspects of experimental studies. Each experiment including human subjects must be approved by Ethical Committees.

A cornerstone document of human research ethics is the Declaration of Helsinki which represents a set of ethical principles regarding human experimentation developed for the medical community (30). It was adopted at the 18th World Medical Association General Assembly in Helsinki, Finland in June 1964, as a statement of ethical principles for medical research involving human subjects, including research on identifiable human material and data. The declaration was amended several times, while the latest one was in Seoul, Korea in 2008.

The Declaration of Helsinki is based upon the duty of researcher to protect the life, health, dignity, integrity, right to self-determination, privacy and confidentiality of personal information of human subjects in research. In addition, it is stated that medical research involving human subjects must conform to generally accepted scientific principles, based on a thorough knowledge of the scientific literature, other relevant sources of information, and adequate laboratory and, as appropriate, animal experimentation.

Apart from the Declaration of Helsinki, international guidelines in medical research are published by the CIOMS (Council for International Organization of Medical Sciences) (4,31).

Exercises

Teaching methods for this topic would include distribution of several published papers; in small groups, students will discuss on appropriateness of used design, the validity of study and the authors' conclusions in the light of the stated objectives.

Tasks 1-5 refer to Examples 1-5 presented previously in the module, while Tasks 6-8 are adapted from Biglan, Norell, and Omenn (32-34).

Task 1

This task refers to the Example 1. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. What are characteristics of stratified randomization?
3. What does placebo mean?
4. What does mean double-blinded trial?
5. What was intervention?
6. How were effects of intervention assessed?
7. What do the results obtained suggest and has the objective been achieved?
8. Discuss the ethical considerations of the study.

Task 2

This task refers to the Example 2. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. How were effects of intervention assessed?
3. What do the results obtained suggest and has the objective been achieved?
4. Discuss the ethical considerations of the study.

Task 3

This task refers to the Example 3. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. What was essential in prophylactic trials?
3. What was intervention?
4. How were effects of intervention assessed?

Task 4

This task refers to the Example 4. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. What was essential in field trials?
3. Why women aged 40-60 were included in this study?
4. What was intervention?
5. How were effects of intervention assessed?

Task 5

This task refers to the Example 5. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. What was essential in community trials?
3. Why were women aged 40-60 years included in this study?
4. What was intervention?
5. How were effects of intervention assessed?
6. What could be another way of controlling an experiment, apart from measuring the dental health status in similar but untreated low-fluoride community?

Task 6

Carefully consider the paper entitled "A randomized controlled trial of a community intervention to prevent adolescent tobacco use" by Biglan et al. (32). Discuss the following questions:

1. What was the objective of this study?
2. Describe and discuss study design.
3. Which type of randomization was used?
4. What was intervention?
5. Were objective outcome criteria developed and used?
6. How were effects of intervention assessed?
7. What do the results obtained suggest and has the objective been achieved?
8. What were the limitations of this study?

Task 7

Mass screening with mammography for early detection and treatment of breast cancer in women can reduce the risk of advanced stages of the disease and death. To investigate this, a randomized trial was conducted among 162,981 women age 40 or more and living in two counties in Sweden at the time of randomization. Each county was divided into 19 blocks selected to give relative socioeconomic homogeneity within each block. In one of the two counties, each block was divided into two units of roughly equal size. One of these units was selected randomly to receive, and the other not to receive, the screening program (33). Discuss the following questions:

1. What was the unit of randomization?
2. What were the advantages and disadvantages of this approach?

Task 8

Read carefully the article "Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease" by Omenn et al. (34). Try to find answers to the following list of questions:

1. Specify the main hypothesis, main outcome and main exposure.
2. Discuss appropriateness of the study design.
3. Rephrase could any other study design be used for this research?
4. What was the target population?
5. What was the main result?
6. Was adjustment for potential confounders carried out?

Assessment of students

Here are some examples of multiple-choice questions for the assessment of students:

1. Which of the following studies does NOT belong to observational studies?

a) cross-sectional study
b) case-control study
c) intervention study
d) prospective cohort study
e) ecological study

2. The major purpose of randomization in an intervention study is to:

a) facilitate double-masking
b) reduce selection bias
c) reduce information bias
d) facilitate measurement of outcome variables
e) avoid sampling variation

3. The problems pertaining to intervention studies include the following items, except one:

a) ethical considerations
b) response and attrition problems
c) high likelihood of comparability of study groups
d) limited feasibility
e) high expenses

4. Comparison of experimental group to a control group could be compared to:

a) placebo
b) another intervention
c) same intervention at a higher dose (or longer duration)
d) no intervention
e) all of the above

5. Clinical trial:

a) deals with subjects who are free of disease but presumed to be at risk
b) is an analysis of risk factors and follows a group of people who do not have the disease
c) involves a population as a whole in which a preventive or therapeutic regimen is allocated
d) is an experiment performed on human beings in order to evaluate a certain therapy
e) is a study in which subjects are not randomized to the exposed or unexposed groups

6. Field trial:

a) is a study in which subjects are not randomized to the exposed or unexposed groups
b) involves a population as a whole in which a preventive or therapeutic regimen is allocated
c) is an experiment performed on human beings in order to evaluate a certain therapy
d) is an analysis of risk factors and follows a group of people who do not have the disease
e) deals with subjects who are free of disease but presumed to be at risk

7. Community trial:

a) is an experiment performed on human beings in order to evaluate a certain therapy
b) involves a population as a whole in which a preventive or therapeutic regimen is allocated
c) deals with subjects who are free of disease but presumed to be at risk
d) is a study in which subjects are not randomized to the exposed or unexposed groups
e) is an analysis of risk factors and follows a group of people who do not have the disease

References

1. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
2. Last JM, editor. A dictionary of epidemiology. Fourth edition. New York, Oxford, Toronto: Oxford University Press; 2001.
3. TheFreeDictionary's Medical dictionary [homepage on the Internet]. Available from URL: http://medical-dictionary.thefreedictionary.com/. Accessed June 5, 2013.
4. Beaglehole R, Bonita R, Kjellstrom T. Basic Epidemiology. Geneva: World Health Organization; 2007.
5. MedicineNet.com. MedTerms dictionary [homepage on the Internet]. Available from URL: http://www.medicinenet.com/script/main/hp.asp. Accessed June 5, 2013.
6. Aschengrau A, Seage GR III. Essentials of epidemiology in public health. Sudbury, MA: Jones&Bartlett Publishers; 2003.
7. Dos Santos Silva I. Cancer epidemiology. principles and Methods. Lyon: IARC; 1999.
8. Lilienfeld DE, Stolley PD. Foundation of epidemiology. New York, Oxford: Oxford University Press; 1994.
9. Hennekens CH, Buring JE. Epidemiology in medicine. Boston/Toronto: Little, Brown and Company; 1987.
10. Gordis L. Epidemiology. Philadelphia: W.B. Sounders Company; 2000.
11. Farmer R, Miller D, Lawrenson R. Epidemiology and public health medicine. Oxford: Blackwell Science; 1996.
12. Singleton RA Jr, Straits BC. Approaches to social research. Third edition. New York: Oxford University Press; 1999.
13. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. London: Arnold Publishers; 2000.
14. Spriet A, Dupin-Spriet T, Simon P. Methodology of clinical drug trials. Basel: Karger; 1994.
15. U.S. National Institutes of Health, National Cancer Institute. Dictionary of cancer terms [homepage on the Internet]. Available from URL: http://www.cancer.gov/dictionary/. Accessed June 5, 2013.

16. Reeves BC. Principles of research: limitations of non-randomized studies. Surgery. 2008;26:120-4.
17. U.S. National Institutes of Health, Clinical Trials.gov. Glossary of clinical trials terms [homepage on the Internet]. Available from URL: http://clinicaltrials.gov/ct2/info/glossary#ttrials. Accessed June 5, 2013.
18. European Union. Biostatistical methodology in clinical trials. Available from URL: http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/vol-3/pdfs-en/3cc7aen.pdf. Accessed June 5, 2013.
19. Friedman LM, Schron EB. Methodology of intervention trials in individuals. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. Oxford: Oxford University Press; 2002. p.569-582.
20. Franceschi S, Plummer M. Intervention trials. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.345-370.
21. U.S. National Institutes of Health, Clinical Trials.gov. Understanding clinical trials [homepage on the Internet]. Available from URL: http://clinicaltrials.gov/ct2/info/understand#Q18. Accessed June 5, 2013.
22. University of Pittsburgh, Research Coordinator Orientation. Comparison of clinical trial phases. Pittsburgh: University of Pittsburgh, Research Coordinator Orientation; 2002. Available from URL: http://www.clinicalresearch.pitt.edu/docs/comparison_of_clinical_trial_phases.pdf. Accessed June 5, 2013.
23. Fischl MA, Richman DD, Hansen N, Collier AC, Carey JT, Para MF, et al. The safety and efficacy of Zidovudine (AZT) in the treatment of subjects with mildly symptomatic human immunodeficiency virus type 1 (HIV) infection: A double-blind, placebo-controlled trial. Ann Inter Med. 1990;112:727-37.
24. Volberding PA, Lagakos SW, Koch MA, Pettinelli C, Myers MW, Booth DK, et al. Zidovudine in asymptomatic human immunodeficiency virus infection: a controlled trial in persons with fewer than 500 CD4-positive cells per cubic millimeter. N Engl J Med. 1990;322:941-9.
25. Ngugi EN, Plummer FA, Simonsen JN, Cameron DW, Bosire M, Waiyaki P, et al. Prevention of transmission of human immunodeficiency virus in Africa: effectiveness of condom promotion and health education among prostitutes. 1988;332:887-90.
26. Shapiro S. Evidence on screening for breast cancer from a randomized trial. Cancer. 1977;39:2772-82.
27. Ast DB, Finn SB, McCaffery I. The Newburgh-Kingston Caries Fluorine Study. Dental findings after three years of water fluoridation. Am J Publ Health. 1950;40:716-24. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1528792. Accessed June 5, 2013.
28. Hoffmeister H, Mensik GBM. Community-based intervention trials in developed countries. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. Oxford: Oxford University Press; 2002. p.583-597.
29. Hill AB. Short Textbook of Medical Statistics. Philadelphia: JB Lippincott Company; 1977.
30. World Medical Association. World Medical Association Declaration of Helsinki: Recommendations Guiding Medical Doctors in Biomedical Research Involving Human Subjects. Available from URL: www.who.int/bulletin/archives/79(4)373.pdf. Accessed: June 5, 2013.
31. Council for International Organizations of Medical Sciences (CIOMS). International Ethical Guidelines for Biomedical Research Involving Human Subjects. GENEVA: CIOMS, 2002. Available from URL: www.cioms.ch/publications/layout_guide2002.pdf. Accessed June 5, 2013.
32. Biglan A, Ary DV, Smolkowski K, Duncan T, Black C. A randomized controlled trial of a community intervention to prevent adolescent tobacco use. Tob Control. 2000;9:24-32.
33. Norell SE. Workbook of Epidemiology. New York, Oxford: Oxford University Press; 1995.
34. Omenn GS, Goodman GE, Thornquist MD. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. New Engl J Med. 1996;334:1150-5.

Recommended readings

1. Dos Santos Silva I. Cancer epidemiology. principles and Methods. Lyon: IARC; 1999.
2. Olsen J, Saracci R, Trichopoulos D. Teaching epidemiology: a guide for teachers in epidemiology, public health and clinical medicine. New York, Oxford, Toronto: Oxford University Press; 2001.
3. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
4. Vetter N, Matthews I. Epidemiology and public health medicine. Edinburgh, London, New York: Churchill Livingstone; 1999.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Public health intervention programmes and their evaluation**
Module: 2.17	ECTS (suggested): 0.1
Author(s), degrees, institution(s)	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Jozica Maucec Zakotnik, MD, Institute of Public Health of the Republic of Slovenia, Ljubljana, Slovenia Zlatko Fras, MD, PhD, Assistant Professor University Clinical Center Ljubljana, Department of Vascular Diseases, Slovenia
Address for correspondence	Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Arterial hypertension, blood pressure, CINDI, evaluation, public health intervention programmes.
Learning objectives	After completing this module students should: • increase their knowledge about evaluation process; • understand and differentiate between different types of evaluation; • understand the public health programme process; • understand the meaning and importance of the last step in this process – the evaluation of public health programmes; • be able to critically assess the limits of evaluation process of public health programmes.
Abstract	A public health intervention is an intervention which is applied to many, most, or all members of a community, with the aim of delivering a net benefit to the community or population as well as benefits to individuals. Every intervention programme has its cycle. One of the most important phases in this cycle is the evaluation phase. Slovenia as a state officially joined international CINDI programme at the beginning of the 1990s, when its activities were limited to Ljubljana demonstrational area. First few years were used as an introductory period of the programme, while more systematically organized activities begun in the late 1990s. The case study presents the different types of evaluation and, as an illustration, the case of evaluating the effectiveness of the CINDI programme activities in Slovenia for reduction of arterial hypertension.
Teaching methods	Teaching methods include introductory lecture, case study, small group discussions, and the whole group discussion (snowball method). After the introductory lecture, students need to read carefully the suggested paper on the subject. Afterwards, they need to answer the questions and discuss the issue - first in small groups and afterwards in the whole group. Students are especially addressed to critically discuss on the limits and strengths of evaluation of public health programmes.
Specific recommendations for teachers	• work under teacher supervision/individual work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on case problem presentation and oral exam.

PUBLIC HEALTH INTERVENTION PROGRAMMES AND THEIR EVALUATION

Lijana Zaletel-Kragelj, Jozica Maucec Zakotnik, Zlatko Fras

Theoretical background

General definitions

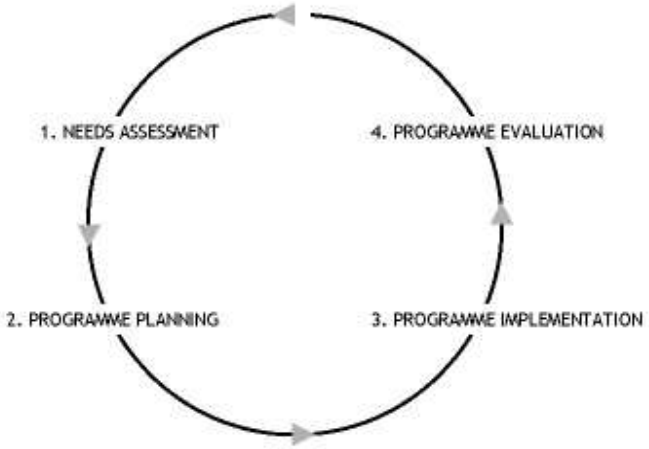
Programme

Several definitions exist on what the term programme means, among which we can find the following:

- a formal set of procedures to conduct an activity (1);
- a set of projects designed to achieve common, long-term goals (2);
- a set of interventions, activities or projects that are typically implemented by several parties over a specified period of time and may cut across sectors, themes and/or geographic areas (3).

Every programme has several steps which could be schematically presented as a cycle (Figure 1), which is to a certain extent similar to the evidence-based public health cycle (4):

Figure 1. The programme cycle



- needs assessment – this is a step at which information about community health problems are gathered. On the basis of this information assumptions on how the needs/problems could be addressed, and the objectives/goals that should be reached are set up;
- programme planning – planning step - uses the assumptions set up at the previous step to plan a programme of activities;
- programme implementation – implementation step - refers to the follow-up of the activities in accordance with the plan. Implementation could be expressed in terms of operational or action plans which commonly outline concrete activities, time frames, responsibilities, budgets etc., for the achievement of different objectives of the programme;
- programme evaluation – by running a programme, we want to know how far the programme went and how effective it is in achieving its goals/objectives. We are able to answer these questions by performing the so-called programme evaluation process. According to World Health Organization (WHO) (2) a programme evaluation is a periodic review and assessment of a programme to determine, in light of current circumstances, the adequacy of its objectives and its design, as well as its intended and unintended results. This process bases on continuous careful monitoring of the course/implementation of the programme.

Programme evaluation is of key importance since on one hand it generates information that can help to improve the programme, and on the other hand it can demonstrate to stakeholders (e.g. funders) and others the impact and the efficiency of the programme.

Intervention and public health intervention

Several definitions exist on what an intervention is, among which we can identify the following:

- an action or entity that is introduced into a system to achieve some results. In the program evaluation context, an intervention refers to an activity, project or program that is introduced or changed (amended, expanded, etc.) (3);
- an action or programme that aims to bring about identifiable outcomes (5).

Planned/desired effects of an intervention expressed in terms of outcomes are general objectives of an intervention.

A public health intervention is an intervention which is applied to many, most, or all members of a community, with the aim to deliver a specific benefit to the community or population as well as benefits to individuals (5,6). Public health interventions include:

- policies of governments and non-governmental organisations;
- laws and regulations;
- organisational development;
- community development;
- education of individuals and communities;
- engineering and technical developments;
- service development and delivery; and
- communication (including social marketing).

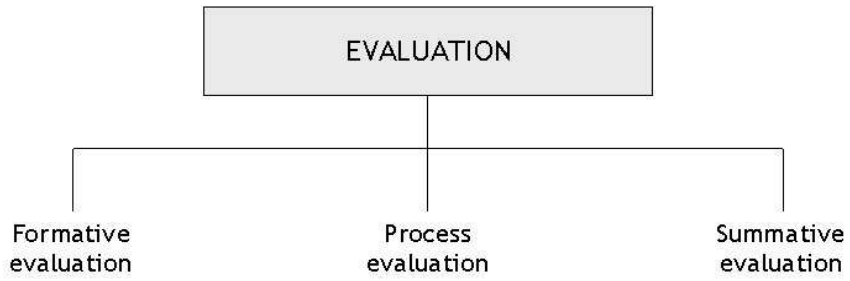
Evaluation

To evaluate something it means literally to look at, and judge, its quality or value. Several formal definitions of evaluation exist, two of them being the following:

- a process that attempts to determine as systematically and objectively as possible the effectiveness and impact of activities according to their objectives (1);
- a systematic and objective assessment of an on-going or completed project, program or policy (3).

Whatever we evaluate, it can be assessed from several different points of view. The three main types of evaluation are formative evaluation, process evaluation, and summative evaluation (Figure 2) (2,3,5,7,8).

Figure 2. The three main types of evaluation



1. **Formative evaluation**
   Formative evaluation is the process of testing programme plans, messages, materials, strategies, and activities for feasibility, appropriateness, acceptability, and applicability to the programme and the target population.
   This type of evaluation is generally used when a new programme is being developed or an existing programme is being modified. It is a process which is often going on during the phase of planning and the course of the programme. In this case it is an evaluation concerned with examining ways of improving and enhancing the implementation and management of an intervention.
   Formative evaluation is aimed at being conducted to help those managing the intervention with the intention of improving their work.
2. **Process evaluation**
   Process evaluation is an assessment of the process of programme delivery. It is dealing with documenting the intensity of interventions, their outreach costs, their short- and medium-term impacts.
   The primary aim of this type of evaluation is to provide practical feedback to those who are responsible for an intervention, so that they can improve its design and performance during the course of the programme.
3. **Summative evaluation**
   Summative evaluation is sometimes referred to as programme impact, or outcome evaluation. Some authors even treat outcome and impact evaluation as two different types of evaluation.
   This type of evaluation is concerned with determining the essential effectiveness of a programme. It is used to determine how well the programme achieved the goal (for example the reduction of morbidity or mortality).
   For this type of evaluation the baseline data must be collected prior to the programme implementation and following the implementation of the programme to determine programme effects. Documenting changes in morbidity and mortality requires a large study population, as well as the analysis of the same type of data for a similar population that did not receive the programme intervention (control group).

Summative evaluation is aimed at being conducted to help external actors (groups who are not directly involved in the management of a programme), for reasons of accountability, or to assist in the allocation of budgetary resources.

Community-based intervention projects/programmes

Many different community-based intervention projects/programmes were designed and/or implemented since the early 1970s to combat chronic non-communicable diseases. A great number of them have aimed at promoting risk-reduction lifestyle changes in different populations. Most of these projects started in the field of cardiovascular disease prevention and emphasized the fact that merely providing risk-reduction measures for people at high risk in health service settings would have only a limited impact to the broader society, e.g. the whole nation/country. On the other hand, if the population as a whole is to be targeted, even a modest change in risk factor and healthy-heart lifestyle would potentially have a huge public health impact. One of the first among such community-based heart health intervention projects was the North Karelia Project (in Finland) which started in 1972 (9).

Many of intervention programmes became international. Countrywide Integrated Non-communicable Diseases Intervention programme (CINDI), which started to spread its ideas in the 1980s, is one of them (10,11).

Case study

Evaluation of the effectiveness of the WHO CINDI programme in Slovenia in combating arterial hypertension

The CINDI recognized the arterial hypertension control as an important element of controlling the overall risk of cardiovascular diseases since arterial hypertension is one of the most important modifiable risk factors for cardiovascular diseases, and among major contributors to the mortality and/or disability of adult population in many countries (12). Changing the lifestyle of the population, screening for arterial hypertension, and early lifestyle and antihypertensive drug treatment are among the cornerstones of the successful prevention of cardiovascular diseases (12-14).

CINDI Working Group on Hypertension worked out international recommendations, where the most frequently used non-pharmacological interventions for arterial hypertension control are: stress management, smoking cessation, salt, calories and saturated fats intake reduction, increase in vegetable/fruit consumption and regular physical activity, and alcohol intake reduction (15). Positive experiences from Finland (North Karelia) and Lithuania showed that the CINDI interventions could be extremely successful in tackling the problem of hypertension (16).

Needs assessment and implementation of an intervention programme in Slovenia

Combating non-communicable diseases has become one of the most important contemporary public health issues in Slovenia since mortality is still mainly attributable to non-communicable diseases, with cardiovascular diseases as the leading group among the causes of death, both in women and men (17).

Among biological/physiological risk factors for non-communicable diseases in Slovenia, more precisely its central part, i.e. Ljubljana area, the high prevalence of severe arterial hypertension (systolic/diastolic blood pressure ≥160/95 mmHg) was registered at the beginning of the 1990s, being 19% (18). This result classified Slovenia among countries with the highest prevalence of severe arterial hypertension (Hungary: 5%, Israel: 7%, Romania 5-10%, Italy: 24-27%) (18).

Once the problem was realized, it was also acknowledged that a programme like CINDI was strongly needed in Slovenia. The process started in the early nineties. First years were introductory and the programme was limited to demonstrational Ljubljana region.

After the second CINDI survey, health promotion and cardiovascular disease prevention philosophy started to spread countrywide. This process outgrow into the Nationwide Programme on Primary Prevention of Cardiovascular Diseases, launched under the auspices of the Ministry of Health of the Republic of Slovenia in autumn 2001 and legally introduced and for the first time carried out at the beginning of 2002 (19,20). Tackling the problem of high prevalence of arterial hypertension was one of the priorities.

Evaluation of an intervention programme

What exactly was done in the observed period to increase the control over the problem of arterial hypertension in Slovenia and could be assigned to CINDI programme? To get this answer, the quantitative and qualitative analyses were performed.

Quantitative evaluation

In 2006, an important study was carried out to evaluate the CINDI Programme contribution in controlling arterial hypertension in Slovenia. It was a study of changes in blood pressure over time (21). In this study, the hypothesis was confirmed that the average values of blood pressure in Ljubljana area after being adjusted for the effects of gender, age and education level globally increased or remained stable in the first half of 12-year period, while they decreased in the second half. In general, the most important finding was the prominent decrease in the average value of systolic blood pressure in the period 1996/97-2002/03. At the same time, the increase of the

diastolic blood pressure between 1990/91 and 1996/97, and its decrease between 1996/97 and 2002/03 were registered (both changes were nearly statistically significant).

The changes in the arterial hypertension prevalence were characterized by its prominent increase between 1990/91 and 1996/97, while during the period of 1996/97-2002/03 a minor decrease was registered. However, this decrease was still important in an epidemiological sense since the trend reverted from an increasing to a decreasing one. These results suggested some unfavourable influence on the blood pressure of Ljubljana area adults between 1990/91 and 1996/97, while within the period 1996/97-2002/03 the influence was favourable. When commenting the observed blood pressure dynamics it was obvious that the favourable changes in blood pressure happened after implementation of the CINDI Slovenia Programme but, on the other hand, it was necessary to take into account also an obvious fact that the 1990/91-1996/97 period was characterized by a very intensive political and socio-economic transition after Slovenia had become an independent state in 1991. This perturbation resulted in important changes in the lifestyle of the population (22). It is very likely that those changes were reflected also in blood pressure levels as well as in many other cardiovascular risk factors (23,24). Because of this fact, the qualitative analysis was necessary to be performed to supplement the quantitative analysis.

Along with this study, other studies were performed with similar findings (25,26), as well as other studies which tried to evaluate the programme from a different standpoint (27). From this perspective, the results of evaluation of the multi-sectorial and multi-disciplinary project Mura, which started in 2001 in Pomurje region, are also very important (25,26). The studies showed that public health interventions in only a couple of years offered several extremely positive results (25). The Mura project is an ongoing intervention project based on the pattern of a process similar to the project carried out in Finland, which proved highly successful and effective (9). Numerous multi-sectorial activities, including primary health care prevention activities, were focused on changing the nutritional and physical activity behaviour of the population and have been in process since the end of 2001 at the regional (first in the Pomurje region), as well as at the national level (25,28). With regard to health prevention activities, specific socio-economic and cultural circumstances were taken into consideration. At the level of population groups at high risk, the concrete health promotion and health education approach was already applied in Beltinci Community in the Pomurje region (25) where the prevalence of many unhealthy behaviours were found as being the highest in Slovenia (29), as well as a combination of multiple risky behaviours (30). For example, according to the first analysis of the CINDI Health Monitor survey (31) for the year 2004, only the prevalence of every-day consumption of sweet soft drinks decreased from 42.9% in 2001 to 29.1% in 2004 (26). The same study showed a strong shift to more healthy behaviours also in use of fat for food preparation. The percentage of people using lard decreased from 30.3% in 2001 to 20.8% in 2004, while the percent of people using olive oil increased from 7.1% to 15.2%. Results of another study, also based on CINDI Health Monitor and CINDI Risk Factor methodology (32), showed considerable improvement not only in healthy behaviours, but also in some of the physiological risk factors. For example, only one year after the start of the programme in the Beltinci community, the average total cholesterol value decreased by 4.9% (25). This programme was already spread from this community to other areas of Slovenia as part of the implementation of the already mentioned nation-wide strategy for prevention of cardiovascular diseases (28). The results are very promising and stimulating for people personally interested and motivated, but sustainability is still under question.

Qualitative evaluation

We have just emphasized that the 1990/91-1996/97 period in Slovenia was characterized by a very intensive political and socio-economic transition with all accompanying problems, potentially also worsening the health status of the population as a whole. Unfortunately, in the period 1990/91-1996/97 the population approach to control cardiovascular risk factors was not among the priorities in the field of Slovene public health.

On the contrary, Slovenia approached the very intensive implementation of the CINDI programme principles during the second half of the 1990s. Moreover, the 1996/97-2002/03 period was characterized by some prominent achievements in the context of spreading CINDI philosophy out of CINDI Slovenia Preventive Unit. Such a situation represented also an obvious need for the broader (national/countrywide) support and implementation. As a consequence, the national alliance, Slovene National Forum on cardiovascular diseases Prevention, was established in 1999 (founded by the Slovene Society of Cardiology) (19), which brought together all important stakeholders in the field of cardiovascular diseases prevention in Slovenia (health care performers, public health institutions, governmental representatives, as well as many relevant scientific, medico-professional, and non-governmental societies) (19). At that time, a joint collaboration group, initiated by the CINDI Slovenia Preventive Unit, elaborated the project on national cardiovascular diseases and other non-communicable diseases prevention programme. The most important interventional arm of the project was the provision of the health counselling and education for individuals at high risk for cardiovascular diseases, which is being performed within the network of 60 Health Education Centres, established all over Slovenia, and coordinated by the CINDI Slovenia Preventive Unit (33). Based on this project, the Ministry of Health decided to support the implementation of the Nationwide Slovene Programme on Primary Cardiovascular Diseases Prevention, the main characteristics of which were already described above. The necessary financial resources were provided by the National Health Insurance Institute of Slovenia (19,20). Within this project, all practicing general practitioners participate in screening the adult population in certain age groups. Its intermediate aims were to decrease the cardiovascular diseases risk factors prevalence

(where arterial hypertension is among the most important) in the population, mainly by lifestyle changes, and to improve early detection and treatment of major cardiovascular risk factors (19,20).

CINDI programme represented the most important base of health promotion and cardiovascular diseases prevention activities described above, inducing the majority of activities towards reducing the arterial hypertension prevalence in Slovene population. These activities took place in various institutions, while CINDI Slovenia Preventive Unit group by itself was especially intensively directed to activities for salt reduction in food industry, and enhancement of healthy behaviours (e.g. nutrition patterns and physical activity). It was also involved in the CINDI-EuroPharm-Forum project, which aimed at enhancing the role of the pharmacists in blood pressure management. We could justifiably claim that without the activities of the CINDI Slovenia programme, the gap in blood pressure control between various population groups could be even larger as it currently is.

Exercises

Task 1

Read carefully the following paper:

Bulc M, Fras Z, Zaletel-Kragelj L. Twelve-year Blood Pressure Dynamics in Adults in Ljubljana Area, Slovenia: Contribution of WHO Countrywide Integrated Non-communicable Diseases Intervention Program. Croat Med J 2006;47:469-77. Available from: URL: http://www.cmj.hr/2006/47/3/16758526.pdf.

Task 2

Try to classify the evaluation presented in this study in appropriate type of programme evaluation.

Task 3

Discuss the process presented in the paper with other students.

Task 4

In bibliographic database (e.g. PUBMED or PUBMED CENTRAL) try to find another paper on this subject.

Task 5

Repeat the exercises No.2 and No.3.

Task 6

Critically discuss strengths and limitations of evaluation of public health intervention programmes with the whole group of students.

Acknowledgements

The authors thank Mateja Bulc, MD PhD, for her contribution to this module.

References

1. Last JM. A dictionary of epidemiology. Oxford: Oxford University Press, 2001.
2. World Health Organization, Regional Office for Europe. Handbook for process evaluation in noncommunicable disease prevention. CINDI 2000. Copenhagen: World Health Organization, Regional Office for Europe, 1999.
3. Planning and Performance Management Unit. Office of the Director of U.S. Foreign Assistance. Glossary of Evaluation Terms. Washington D.C.: Planning and Performance Management Unit. Office of the Director of U.S. Foreign Assistance, 2009.
4. Brownson RC, Baker EA, Leet T, Gillespie KN. Evidence based public health. Oxford, New York: Oxford University Press, 2003.
5. Rychetnik L, Hawe P, Waters E, Barratt A, Frommer M. A glossary for evidence based public health. J Epidemiol Community Health 2004;58;538-545.
6. McLeroy K, Norton B, Kegler M, et al. Community-based intervention. Am J Public Health 2003;93:529-33
7. Clarke L (ed.). Handbook on Monitoring and Evaluation for the CINDI Network. Pietermaritzburg: Children in Distress CINDI Networking, 2006.
8. Mallonee S. Evaluating Injury Prevention Programs: The Oklahoma City Smoke Alarm Project. http://www.futureofchildren.org. Accessed: May 29, 2013).
9. Puska P, Tuomilehto J. Nissinen A, Vartainen E. The North Karelia project. 20 year results and experiences. Helsinki, National Public Health Institute; 2001.
10. World Health Organization, Regional Office for Europe. CINDI Countrywide Integrated Non-communicable Diseases Intervention Programme. Protocol and guidelines for monitoring and evaluation procedures. Berlin: Springer-Verlag; 1987.

11. World Health Organization, Regional Office for Europe. A strategy to prevent chronic disease in Europe. A focus on public health action. The CINDI vision. Copenhagen: WHO, Regional Office for Europe; 2004.
12. Nissinen A, Berrios X, Puska P. Community-based noncommunicable disease interventions: lessons from developed countries for developing ones. Bull World Health Organ 2001;79:963-70.
13. Yach D, Hawkes C, Gould CL, Hofman KJ. The global burden of chronic diseases: overcoming impediments to prevention and control. JAMA 2004;291:2616-22.
14. Williams B, Poulter NR, Brown MJ, Davis M, McInnes GT, Potter JF, et al. British Hypertension Society guidelines for hypertension management 2004 (BHS-IV): summary. BMJ 2004;328:634-40.
15. World Health Organization, Regional Office for Europe. High blood pressure management. Report on a WHO training seminar. Barcelona, Spain, 24–26 October 1996. Copenhagen: WHO, Regional Office for Europe; 1996.
16. Glasunov IS, Grabauskas V, Holland WW, Epstein FH. An integrated programme for the prevention and control of noncommunicable diseases. A Kaunas report. J Chron Dis 1983;36:419-26.
17. World Health Organization, Regional Office for Europe. The Health for All Statistical Database. Copenhagen: WHO, Regional Office for Europe; 2005.
18. Wilhelmsen L, Strasser T, Jezersek P, Accetto R, Cibic B, Dobovisek J, et al. WHO-WHL Hypertension Management Audit Project. J Hum Hypertens 1993;7:257-63.
19. Fras Z. Integrality of primary prevention of cardiovascular diseases and its implementation in Slovenia (in Slovene). Slov Kardiol 2004;1:8-23.
20. Fras Z, Maucec Zakotnik J. Life-style interventions in primary health care - implementation of Slovenian CINDI approach as a national strategy for prevention of coronary heart disease. In: International conference Promoting health through physical activity and nutrition, Radenci, 18- 21 April 2002. Final programme, abstracts and papers. Ljubljana: CINDI Slovenia; 2002. p.87-93.
21. Bulc M, Fras Z, Zaletel-Kragelj L. Twelve-year Blood Pressure Dynamics in Adults in Ljubljana Area, Slovenia: Contribution of WHO Countrywide Integrated Non-communicable Diseases Intervention Program. Croat Med J 2006;47:469-77. Available from: URL: http://www.cmj.hr/2006/47/3/16758526.pdf. Accessed: May 29, 2013.
22. Hofmarcher MM. Is Public Health Between East and West? Analysis of Wealth, Health and Mortality in Austria, Central and Eastern European Countries, and Croatia Relative to the European Union. Croat Med J 1998;39:241-8.
23. Wolf HK, Tuomilehto J, Kuulasmaa K, Domarkiene S, Cepaitis Z, Molarius A, et al. Blood pressure levels in the 41 populations of the WHO MONICA Project. J Hum Hypertens 1997;11:733-42.
24. Accetto R, Zemva A, Dolenc P. Control of hypertension in the Republic of Slovenia (in Slovene). Zdrav Vestn 2001;70:279-83.
25. Zaletel-Kragelj L, Maucec Zakotnik J, Belovic B. Assessment of health promotion activities in Health region Pomurje. First results of "CINDI Health Monitor Slovenia 2004" (in Slovene). In: Fras Z (ed.) Slovene Forum for cardiovascular diseases prevention 2005: book of articles. Zdruenje kardiologov Slovenije, Ljubljana, 2005. pp 27–32.
26. Zelko-Peterka E, Zaletel-Kragelj L. Local community health enhancing programme assessment. The case of Beltinci community (in Slovene). Slov Kardiol 2005;2:8-12.
27. Fras Z, Bulc M, Maucec-Zakotnik J, Zaletel-Kragelj L. Prevalence of arterial hypertension, its awareness and control in the adult population of the Ljubljana area of Slovenia. Results of WHO's Countrywide Integrated Noncommunicable Diseases Intervention (CINDI) Program Survey 2002/2003. Slov Kardiol 2005;3:106-114.
28. Fras Z, Maucec Zakotnik J. Life-style interventions in primary health care - implementation of Slovenian CINDI approach as a national strategy for prevention of coronary heart disease. In: International conference: promoting health through physical activity and nutrition. Radenci, Slovenia, April 18–21, 2002. Final programme: abstracts and papers. CINDI Slovenija, Ljubljana, 2002. pp 87–93.
29. Zaletel-Kragelj L, Erzen I, Fras Z. Interregional differences in health in Slovenia. II. Estimated prevalence of selected behavioral risk factors for cardiovascular and related diseases. Croat Med J 2004;45:644–650. Available from: URL: http://www.cmj.hr. Accessed: May 29, 2013.
30. Zaletel-Kragelj L, Fras Z. Identification of population groups with multiple hazardous health behaviours for cardiovascular diseases in Slovenia. Slov Kardiol 2005;2:112–117.
31. Prattala R, Helasoja V, Laaksonen M, Laatikainen T, Nikander P, Puska P. CINDI Health Monitor. Proposal for practical guidelines. Helsinki: National Public Health Institute, 2001.

32. World Health Organization, Regional Office for Europe. Protocol and guidelines: Countrywide Integrated NCD Intervention (CINDI) Programme (Revision of 1994). Copenhagen: WHO Regional Office for Europe, 1996.
33. Vrbovsek S, Luznar N, Maucec Zakotnik J (eds.). Together in protecting and strengthening health: what have we achieved in the first eight years? (in Slovene). Ljubljana: Institut za varovanje zdravja Republike Slovenije, 2009.

Recommended readings

1. Brownson RC, Baker EA, Leet T, Gillespie KN. Evidence based public health. Oxford, New York: Oxford University Press, 2003.
2. Bulc M, Fras Z, Zaletel-Kragelj L. Twelve-year Blood Pressure Dynamics in Adults in Ljubljana Area, Slovenia: Contribution of WHO Countrywide Integrated Non-communicable Diseases Intervention Program. Croat Med J 2006;47:469-77.
3. Puska P, Tuomilehto J. Nissinen A, Vartainen E. The North Karelia project. 20 year results and experiences. Helsinki, National Public Health Institute; 2001.
4. World Health Organization, Regional Office for Europe. Handbook for process evaluation in noncommunicable disease prevention. CINDI 2000. Copenhagen: World Health Organization, Regional Office for Europe.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Surveillance**
Module: 2.18	ECTS (suggested): 0.25
Author(s), degrees, institution(s)	Irena Klavs, MD, PhD, Professor National Institute of Public Health
Address for Correspondence	Irena Klavs National Institute of Public Health Trubarjeva 2, 1000 Ljubljana, Slovenia Irena.Klavs@ivz-rs.si
Keywords	Behavioural risk factor surveillance system, population surveillance, sentinel surveillance, surveillance.
Learning objectives	After completing this module students will: • understand the definitions of surveillance (SU); • understand the aims and objectives of SU and uses of SU information; • understand the elements of SU systems and steps involved; • be able to identify the strengths and limitations of SU systems and consider possibilities for improvements.
Abstract	SU is ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and improve health. Monitoring trends is the cornerstone objective of SU systems. Information on characteristics of individuals with health problems permits identification of groups at highest risk of disease. SU information can provide a documentation of the success of an intervention or indicate the need for one. Information from SU systems can contribute to setting of public health priorities, advocacy, health planning and making decisions regarding the allocation of available resources, monitoring and evaluating of public health programmes. Evaluating public health SU systems is to ensure that problems of public health importance are being monitored efficiently and effectively.
Teaching methods	An introductory lecture gives the students an overview of SU (the theory and a case study) and is followed by an individual work (recommended readings, literature search and preparation for presenting an example of published results from a national SU system including critical analysis of strengths and limitations and suggestions for improvements).
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 35% / 65%; • facilities: a lecture room, a computer room; • equipment: LCD projection, computers, access to the Internet and bibliographic data-bases; • training materials: recommended readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire (MCQ); presentation of an example of published results from a national SU system with critical analysis of strengths and limitations and suggestions for improvements.

SURVEILLANCE

Irena Klavs

Theroretical background

Definition of surveillance

In 1968, the 21st World Health Assembly described surveillance as the systematic collection and use of epidemiological information for the planning, implementation, and assessment of disease control; in short, surveillance implied "information for action".

According to Last (1), surveillance is systematic ongoing collection, collation, and analysis of data and the timely dissemination of information to those who need to know so that action can bet taken. He distinguishes surveillance from monitoring by the fact that it is continuous and ongoing, whereas monitoring is intermittent or episodic.

At greater length, the Centers for Diseases Control and Prevention (CDC) in the US defines surveillance as the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and improve health (2).

Surveillance methods were originally developed as part of efforts to control infectious diseases. Later, the basic concepts of surveillance have been applied to all areas of public health (3).

Aims and objectives of surveillance and uses of surveillance information

The ultimate aim of a surveillance system is the application of surveillance information for prevention and control of diseases. Surveillance information is meant to be and should be used for public health action.

In 1988, the Institute of Medicine defined three essential functions of public health that emphasized the central role of surveillance:

- assessment of the health of communities, which depends largely on surveillance;
- policy development based on the "community diagnosis" and prognosis established through surveillance; and
- assurance that necessary services are provided, using surveillance as one measure of the impact of the programmes (4).

CDC (2) identified the following potential uses for the information resulting from public health surveillance systems:

1. guide immediate action for cases of public health importance;
2. measure the burden of a disease (or other health-related events), including changes in related factors, the identification of populations at high risk, and the identification of new or emerging health concerns;
3. monitor trends in the burden of a disease (or other health-related events), including the detection of epidemics (outbreaks) and pandemics;
4. guide the planning, implementation, and evaluation of programmes to prevent and control diseases, injuries or adverse exposures;
5. evaluate public policy;
6. detect changes in health practices and the effects of these changes;
7. prioritize the allocation of health resources;
8. describe the clinical course of diseases; and
9. provide a basis for epidemiologic research.

The list of principal objectives of surveillance provided by Haden and O'Brien (5) indicates similar uses of information generated by surveillance systems:

1. give early warning of changes of incidence;
2. detect outbreaks early;
3. evaluate the effectiveness of interventions;
4. identify at-risk groups; and
5. help set priorities for resource allocation.

Berkelman, Stroup and Buehler (3) list the following purposes of public health surveillance:

1. to define public health priorities;
2. to characterize disease patterns by time, place and person;
3. to detect epidemics;
4. to suggest hypotheses;
5. to identify cases for epidemiological research;

6. to evaluate prevention and control programmes; and
7. to facilitate planning, including projection of future trends and health care needs.

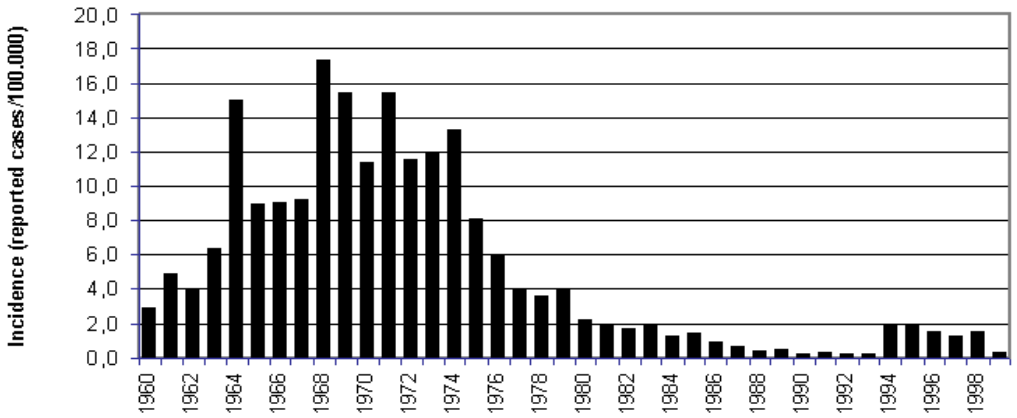
Surveillance systems generally provide descriptive information regarding when and where health problems are occurring and who is affected – the basic epidemiology parameters of time, place, and person.

Monitoring trends

Monitoring trends is the cornerstone objective of most surveillance systems. The detection of an increase in adverse health events can alert public health agencies to the need for further investigation.

For example, after the reported incidence rates of early syphilis in Slovenia, it gradually declined through 1970s and 1980s, whereas in 1994, an 18-fold increase in the annually reported incidence rate of early syphilis (to 1.81 cases per 100,000 population) in comparison to the preceding year was observed (Figure 1). Consequently, the epidemiologist from the Communicable Diseases Centre at the Institute of Public Health of the Republic of Slovenia together with the dermatovenerologist working at the Central Dermatovenerological Clinic in Ljubljana started an investigation. They reviewed the medical records of all notified syphilis cases in 1994. Available information included information on sexual partners, country of probable source of infection, and patients' occupation. The results of this initial investigation indicated that a majority of cases was acquired abroad, in the Russian Federation or in the newly independent states (NIS).

Figure 1. Reported incidence rates of early syphilis, Slovenia, 1960-1999



The annually reported incidence rates remained elevated for a period of five years. In 1999, the surveillance data were analysed for the whole period 1994-1998 (6). More than half of all reported cases (62%) during this period were directly or indirectly linked to a source of infection abroad. Among these, the majority of cases (73%) were linked to the Russian Federation or NIS, where a major syphilis epidemic was evolving. Of these, 68% cases occurred in males, with a high proportion of long-distance lorry drivers.

Detection of outbreaks through monitoring trends in reported incidence rates of diseases is often cited as one of potential uses of surveillance information. In practice, however, very often outbreaks are first detected and reported by astute clinicians, before information on case reports is received and analyzed in a public health agency.

Identifying populations at high risk

Information on characteristics of individuals with health problems permits identification of groups at the highest risk of disease. For example, in 2007, 36 cases of new human immunodeficiency virus (HIV) infection diagnoses were notified in Slovenia. Among these 36 cases, 34 were men and 29 of these men were known to be men who have sex with men (MSM). This information clearly indicated that MSM were the most affected population group in Slovenia.

Information on specific exposures and behaviours may provide insight into the aetiology and modes of transmission. Thus, surveillance information can guide prevention activities even before the aetiology of a disease is known. For example, in the early 1980s, surveillance of the acquired immunodeficiency syndrome (AIDS) provided information on the behavioural characteristics as well as medical histories of AIDS patients. Nearly all had an identified sexual, injecting drug-use or transfusion exposure. This information was sufficient to calm the public that the disease was not transmitted through ordinary social interaction. The surveillance information, together with initial epidemiologic investigations, defined the modes of HIV transmission even before HIV was discovered.

Evaluating interventions

Evaluation of the effect of public health interventions is complex. Full-scale evaluation, for example through randomised, placebo controlled trials or community randomised trials, may not be feasible. Decision makers and health planners often need to make decisions based on the best available information. Trends in the reported incidence rates of a disease identified through surveillance can sometimes provide a convincing documentation of the success of an intervention.

For example, the introduction of childhood vaccination against pertusis in Slovenia in 1959 resulted in a dramatic decline in the reported incidence of pertusis (Figure 2) (7).

Conversely, trends in the reported incidence can indicate a need to introduce an intervention or modify an intervention. In addition, charting the characteristics of affected individuals, surveillance may provide a comparatively inexpensive and sufficient assessment of the impact of an intervention.

For example, despite the high vaccination coverage against pertusis in Slovenia for many decades, the reported incidence has increased slightly after 2002 and further in 2006 to the level of 27.5 per 100,000 population which represented a 6.5-times increase in comparison to the previous year (Figure 2) (7). In addition, marked shift in age distribution among the reported cases was observed with the highest age-specific incidence among 9 and 10 years old in 2006 (Figure 3). This indicated that a booster dose at school entrance or latest at the age of 8 years should be introduced to decrease the transmission of disease among school children and to further reduce the burden of disease among infants.

Figure 2. Reported cases of pertusis, Slovenia, 1957-2006

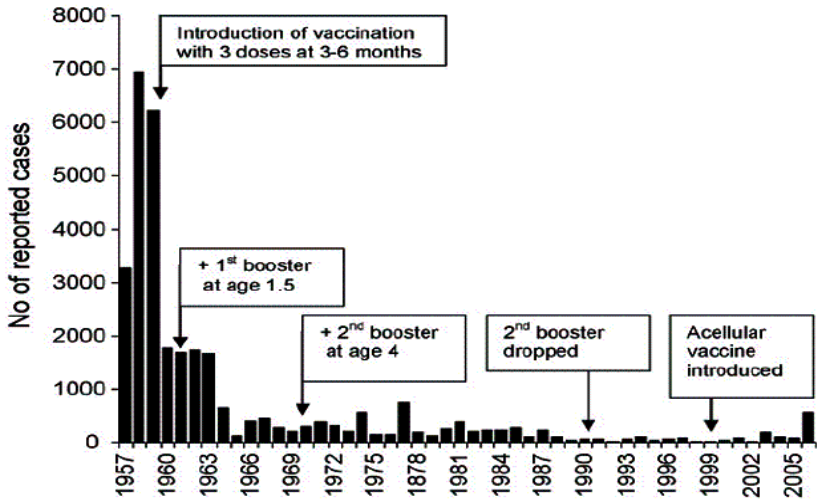
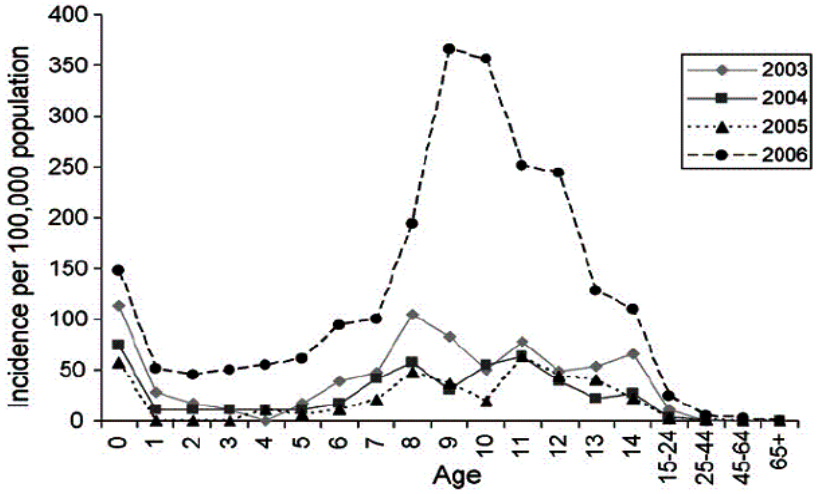


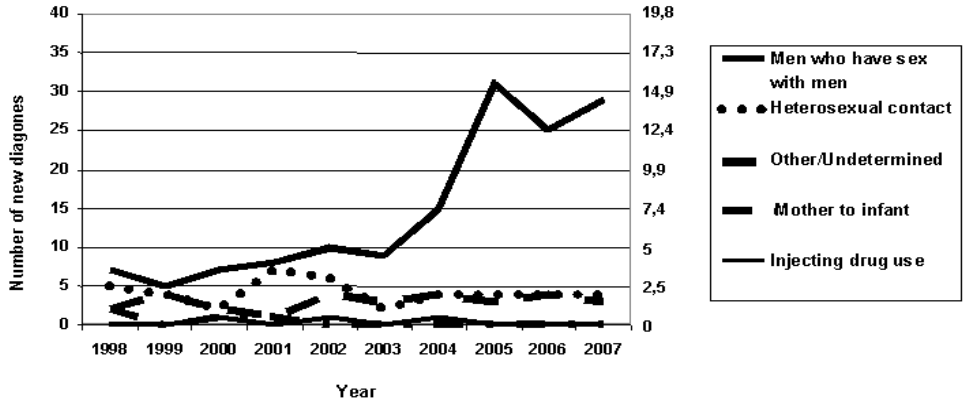
Figure 3. Age specific incidence rates of pertusis, Slovenia, 2003-2006



Setting public health priorities and help in allocating resources

Information from surveillance systems can contribute to setting of public health priorities, advocacy, health planning and making decisions regarding the allocation of available resources, as well as monitoring and evaluating of public health programmes. In addition to political leaders, the information from surveillance systems educates the public, the media, and health care workers directly responsible for providing health care.

Figure 4. Newly diagnosed HIV cases according to transmission, Slovenia, 1998-2007



For example, in Slovenia in 2007, a dramatic increase in reported incidence of newly diagnosed HIV infection cases among MSM was documented (Figure 4). The Institute of Public Health of the Republic of Slovenia at that time presented the results to the Ministry of Health and advocated for the increase in the funding allocated to the prevention of HIV infection and other sexually transmitted infections (STI) among MSM. This resulted in increased funding allocated to selected non-governmental organizations of MSM, which were judged to be best capable to reach this population group with preventive interventions tailored to their specific needs.

Links to services

At the community level, surveillance is often an integral part of the delivery of preventive and therapeutic services by public health institutions. This role is particularly true for infectious diseases where interventions are based on known modes of disease transmission, where therapeutic or prophylactic interventions are available, and thus the receipt of a case report triggers a specific response.

For example, a report of invasive meningococcal disease caused by Neisseria meningitidis, which is spread by respiratory droplets, triggers an epidemiologic investigation with the aim of identifying close contacts at home, in school or elsewhere, who would benefit from post-exposure prophylactic therapy. Thus, at the local level, surveillance information can serve to initiate individual preventive actions.

Elements of surveillance systems, steps involved

Haden and O'Brien (5) identify four steps involved in all surveillance systems:

1. data collection;
2. data collation;
3. data analysis; and
4. data dissemination.

Data collection

There are many potential sources of surveillance data. A few of the major ones used in Slovenia are:

1. mortality data (8);
2. out of hospital care information system (8,9) (includes data on preliminary diagnosis for patients' visits in primary health care and on final diagnoses in the secondary or tertiary level outpatient care);
3. hospital care information system (includes data on hospitalisations and admission diagnoses) (8,10);
4. perinatal information system (11);
5. specific diseases registers (e.g. cancer register (12,13));
6. notifications of communicable diseases (14-19);
7. specific national administrative systems (e.g. information system with data on vaccinations, information system on data about adverse vaccination events) (20,21).

The data collection methods must be defined precisely and understood well by those who collect the data (e.g. clinicians). This prevents misclassification of cases and ensures that more accurate and complete information is collected.

The case definition is fundamental to any surveillance system. It ensures that the same measure for a health-related event under surveillance is used across geographical areas and through time.

For example, in a laboratory surveillance of sexually transmitted Chlamydia trachomatis infection, the data providers in participating laboratories should know well the laboratory surveillance case definition (e.g. at least one of the following three criteria should be met:

- isolation of Chlamydia trachomatis from a specimen of the ano-genital tract or from the conjunctivae;
- demonstration of Chlamydia trachomatis by DFA test in a clinical specimen;
- detection of Chlamydia trachomatis nucleic acid in a clinical specimen).

The desire to collect too detailed information must be tempered by the need to limit data to items that can be readily and consistently collected over long periods of time. For example, in a laboratory surveillance of sexually transmitted Chlamydia trachomatis infection, the data providers in participating laboratories could report the following data items (variables):

- specimen type;
- personal identifier (to eliminate duplicates);
- date of birth (to assist in eliminating duplicates and to compute age at diagnosis);
- gender;
- municipality;
- date of collection of specimen;
- reporting laboratory;
- type of clinical service.

This would provide good enough information to monitor reported rates of new diagnoses of sexually transmitted infection with Chlamydia trachomatis in different age groups of men and women in different municipalities and within different types of clinical services.

In order to identify trends over time, the importance of consistency in the on-going standardised data collection methods (means and data elements) is crucial. To be able to compare surveillance results between different geographical regions or different populations, data collection methods must be standardised.

Data collation

On the national and regional level, surveillance data is often collated in computerised databases, which facilitates routine and more sophisticated analyses and production of regular as well as special surveillance reports. On the local level, it is often not necessary that surveillance data are in a computerized database. The data can be collated in a paper-based system or in a spreadsheet.

Surveillance data may be in the form of individual patient records or aggregate counts and tabulations. Individual cases data permit more flexibility of analysis than aggregated data.

Data analysis

Most often, the analysis of surveillance data does not call for sophisticated data manipulation. Simple descriptive analyses according to basic epidemiological parameters of time, place, and person - showing time trends and distribution of characteristics of cases are usually sufficient for surveillance reports.

A case report may include dates, such as those of the onset of disease, diagnosis, and report to local, regional or national public health authority. Analysis can be based on any of these dates, however, if there are long delays between dates of diagnosis and report, analyses of trends based on dates of diagnosis may be unreliable for the most recent period.

Dissemination

Timely dissemination of information to political leaders, the public, the media, and health care workers directly responsible for providing health care is essential. It is also important that the original providers of data to the surveillance system are given regular feedback. In addition, if resources permit, it is a courtesy to accommodate ad-hoc enquiries from data contributors.

Limitations of surveillance systems

Haden and O'Brien (5) identify four potential shortcomings of surveillance systems, which are similar for all information sources:

1. completeness;
2. accuracy;
3. relevance and/or representativeness; and
4. timeliness.

Completeness (sensitivity)

The importance of completeness or sensitivity of surveillance information depends on the specific objectives of the particular disease surveillance. When surveillance information is used to detect clusters to trigger an intervention and possible outbreak investigation, surveillance systems are quite tolerant to incompleteness. For example, a few reported cases of hepatitis A in a primary school can reasonably be expected to be incomplete, as most cases will be asymptomatic. However, such information is good enough to trigger action to prevent the further spread and an outbreak investigation may ascertain the full number of cases.

If all people with the condition under surveillance in the target population are detected by a surveillance system, then its sensitivity is 100%. For example, the information on reported AIDS cases in Slovenia is assumed to be relatively complete. AIDS is a serious disease, the individual presumably seeks health care and is very likely to be diagnosed correctly and when so, the case is most likely notified. In contrast, HIV infection is asymptomatic for many years and can be diagnosed very late. Thus, the information on the reported newly diagnosed HIV cases can reasonably be expected not to reflect the true incidence of HIV infection.

For comparing surveillance data over time and geographical areas, it is important that the degree of completeness is consistent between all data providers and through time. Otherwise, trends will be distorted and conclusions about the differences or similarities in the disease burden between different geographical regions will be flawed.

Accuracy

Accuracy is most critical for surveillance of diseases with very low incidence, where misdiagnosis or misclassification can generate a pseudo outbreak and trigger inappropriate action.

Representativeness

Representativeness is a measure of how well reported cases in a population reflect all cases that actually occurred in the population. Surveillance reporting is rarely complete, and cases that are reported may differ from those unreported in terms of demographic, behavioural, and risk exposures characteristics, geographic location or use of health-care services.

Timeliness

Surveillance information should be available in a timely manner. The judgement about what is timely may vary according to the health condition under surveillance and its potential consequences for the health of the community. For example, a case invasive meningococcal disease caused by Neisseria meningitidis should be reported immediately, while newly diagnosed cases of breast cancer will not require an immediate public health response and are often reported with longer reporting delays. When very rapid access to surveillance information is important, the system of data collection, the process of data management and the approach to analysis and interpretation should be kept as simple and as easy as possible. Especially when emergency action needs to be taken, for example in a communicable disease outbreak, the real time information based on surveillance data is important. The importance of timeliness may outweigh the need for completeness of surveillance data.

Whatever periodicity is used it should be specified and adhered to by all participants in all phases of the surveillance system loop. Timeliness refers to the entire surveillance cycle, from how quickly cases are reported to the distribution of surveillance reports. With increasing computerization and internet use, reporting at the time of case identification is becoming a reality in some countries.

Evaluating public health surveillance systems

The purpose of evaluating public health surveillance systems is to ensure that problems of public health importance are being monitored efficiently and effectively. In 2001, Centers for Diseases Control and Prevention published the updated guidelines for evaluating public health surveillance systems (2). The following tasks involved are described:

1. engage the stakeholders in the evaluation;
2. describe the surveillance system to be evaluated:
    - describe the public health importance of the health-related event under surveillance,
    - describe the purpose and operation of the surveillance system,
    - describe the resources used to operate the surveillance system;
3. focus the evaluation design;
4. gather credible evidence regarding the performance of the surveillance system:
    - indicate the level of usefulness,
    - describe each system attribute,
        - simplicity,
        - flexibility,
        - data quality,

- acceptability,
- sensitivity,
- predictive value positive,
- representativeness,
- timeliness,
- stability;

5. justify and state conclusions, and make recommendations;
6. ensure use of evaluation findings and share lessons learned.

Conclusion

Effective public health response depends on reliable, continuous flow of information provided by numerous and often complex surveillance systems. Adequate resources needed for surveillance systems and their regular evaluation should be allocated.

Case Study: HIV infection surveillance in Slovenia

Second generation HIV surveillance

HIV infection surveillance information has been crucial in understanding the evolvement of the pandemic and for generating evidence-based global public health response. Information about who is infected and who is at risk of infection can help targeting prevention efforts to slow the spread of HIV. Perhaps the most useful information for targeting HIV prevention is the behavioural surveillance information. Surveillance information also helps assessing and forecasting treatment and care needs for those affected.

The national HIV surveillance system should be tailored to the pattern of the epidemic. According to the World Health Organization (WHO) and the Joint United Nations Programme on HIV/AIDS (UNAIDS), strengthened HIV surveillance systems, dubbed "second generation surveillance systems" aim to concentrate resources where they will yield information that is most useful in reducing the spread of HIV and providing treatment and care to those affected (22-24).

The goals of the second generation HIV surveillance systems are (22):

- better understanding of trends over time;
- better understanding of behaviours driving the epidemic in a country;
- surveillance more focussed on sub-populations at highest risk of infection;
- flexible surveillance that moves with the needs and state of the epidemic; and
- better use of surveillance data to increase understanding and to plan prevention and care.

According to WHO and UNAIDS, approach to national HIV surveillance that will provide most useful surveillance information should result from different data collection mix which should depend on the HIV epidemic state (classified as either low-level, concentrated or generalised) (22).

Low level HIV epidemic is defined as epidemic, where HIV infection may have existed for many years, but has never spread to significant levels in any sub-population and is largely confined to individuals with higher risk behaviour, e.g. sex workers, injecting drug users (IDU), MSM. Numerical proxy is that HIV prevalence has not consistently exceeded five percent in any defined sub-population. In such a low-level HIV epidemic situation, which is the case for Slovenia, the recommended components of the second generation HIV surveillance systems are (22):

- HIV and AIDS case reporting;
- HIV surveillance in sub-populations at risk;
- surveillance of STI and other biological markers of risk;
- tracking of HIV in donated blood; and
- cross-sectional surveys of behaviour in sub-populations with risk behaviour.

Approach to HIV surveillance in Slovenia

HIV infection surveillance information has been crucial in understanding the evolvement of the epidemic in Slovenia and for generating evidence-based public health response.

The different components of the Slovenian national HIV surveillance system are:

1. HIV and AIDS case reporting;
2. monitoring HIV prevalence change in selected sentinel sub-populations at different behavioural risk;
3. collating information about the results of testing of all donated blood units for blood safety purposes;
4. behavioural surveillance in two sentinel sub-populations at highest behavioural risk, MSM and IDU.

We also use the information generated from the STI surveillance system (19).

HIV surveillance information is also complemented by the results of epidemiological studies that are relevant for HIV epidemiology.

HIV and AIDS case reporting

Methods

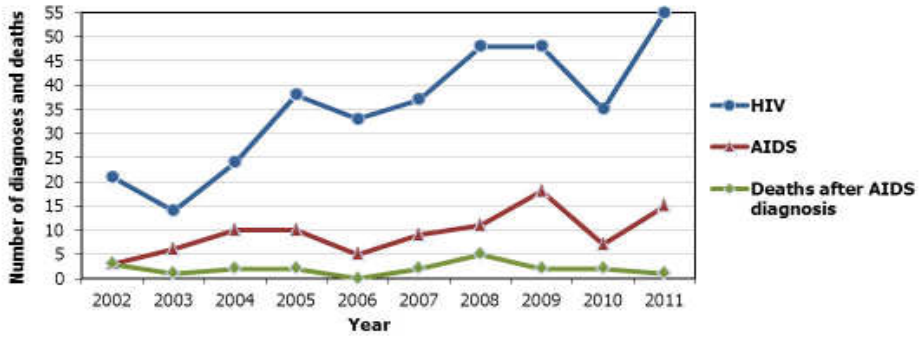
In Slovenia, notification of HIV infection and AIDS is mandatory according to the Communicable Diseases Law (14) and the Law on Health Care Data Bases (16).

We regularly collect, analyse, interpret and publish information about mandatory reported cases of newly diagnosed HIV infections, AIDS cases and deaths in patients diagnosed with AIDS. We use the European AIDS case definition (25). To be able to better interpret this information we also monitor overall national HIV diagnostic testing trends with annual collection of information about the number of HIV tests performed in laboratories.

Results

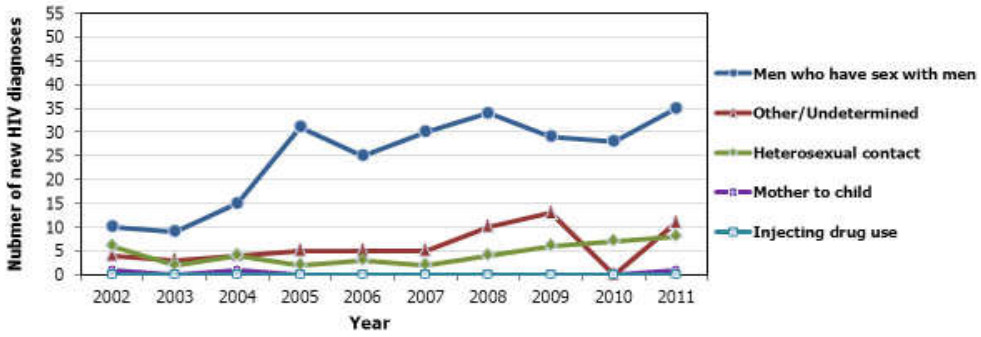
During the past ten years (2002-2011), the reported incidence rates of newly diagnosed HIV cases varied from the lowest 7.0 per million population in 2003 (14 cases) to the highest 26.8 per million population in 2011 (55 cases) (Figure 5). This is rather low in comparison to most countries in the EU.

Figure 5. Newly diagnosed cases of HIV, AIDS and deaths caused by AIDS, Slovenia, 2002-2011



The dramatic increase in the reported HIV incidence after 2003 was due almost exclusively to the increase in new diagnoses of HIV infection among MSM (Figure 6).

Figure 6. Newly diagnosed cases of HIV according to transmission, Slovenia, 2002-2011



The last new HIV diagnosis among IDU was reported more than 10 years ago (in 2001). In 2011, one new HIV diagnosis in a child born to a mother with HIV infection was reported, after the previous case reported in 2004.

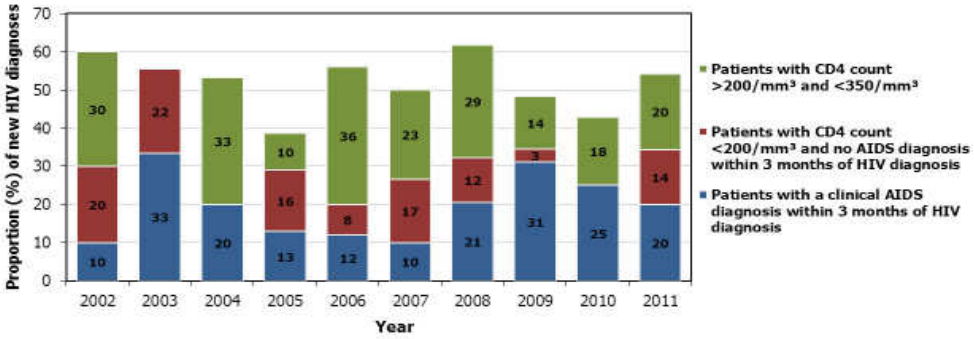
During the past ten years (2002-2011), the annually reported rates of AIDS cases have remained below 10 per million population and deaths among AIDS patients remained below 5 per million population, which also reflects the effect of universal access to highly active anti-retroviral therapy.

For appropriate interpretation of HIV case reporting surveillance results it is important to understand the extent of HIV testing. In general, more extensive HIV testing results in a reduced proportion of undiagnosed HIV infections in the population. Promotion of HIV testing for earlier diagnosis is one of the most important public health interventions. It aims also towards more timely interventions to prevent further transmission among people with newly diagnosed HIV infection.

In Slovenia, overall HIV diagnostic testing rate is relatively low in comparison to most European countries (26), but has been increasing slowly. In 2011, 1.9 diagnostic HIV tests were performed per 100 Slovenians, a three percent increase in comparison to the preceding year.

If HIV infection is diagnosed late, the opportunity for timely and very effective therapy that reduces the risk of early development of AIDS and AIDS related death is missed. Among the MSM, a group with a disproportionally high number of new HIV diagnoses, a substantial proportion of new HIV diagnoses are conducted quite late (at the same time, AIDS is diagnosed either when CD4 cells count has already fallen below 200/mm^3, or when CD4 count is already bellow 350/mm^3) (Figure 7).

Figure 7. Newly diagnosed cases of HIV according to transmission, Slovenia, 2002-2011



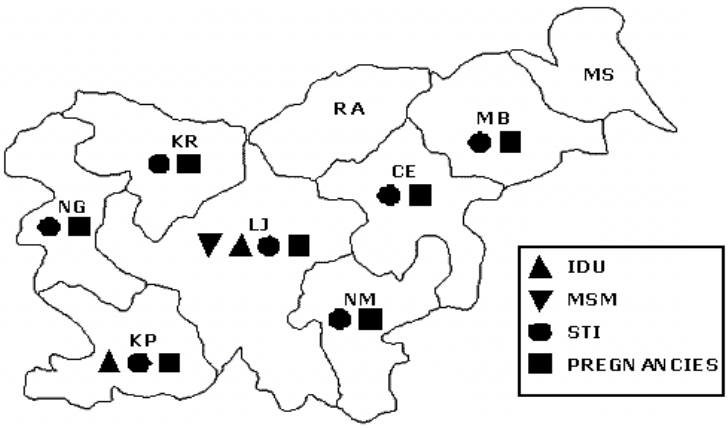
Monitoring HIV prevalence change in sentinel sub-populations

To complement surveillance information on HIV/AIDS case reporting we monitor HIV prevalence with unlinked anonymous HIV testing in several easily accessible sentinel groups at higher behavioural risk (IDU, MSM clients of STI clinics tested for syphilis) and also in one low-risk population group (pregnant women who are screened for syphilis) (27).

Methods

Residual sera from sera specimens obtained from STI patients and pregnant women for syphilis serology are sampled continuously in several laboratories since 1993. Since 1995, saliva specimens are continuously and voluntarily obtained from IDU entering substitution therapy programmes at one or two sites and since 2003 also for two months per year from convenient samples of IDU clients at least one needle exchange harm reduction programme. Voluntary confidential testing for HIV is offered to all IDU included into the sentinel surveillance sample. Since 1996, once per year saliva specimens are voluntarily obtained from a convenient sample of MSM at a community venue in Ljubljana. Safer sex is promoted and information about access to voluntarily confidential and also anonymous HIV testing and counselling is provided to all MSM included into the sentinel surveillance sample. In addition to the information about the type of sentinel population, sampling period and sentinel site, all specimens are labelled only with sex and age group, frozen and stored at -20^0 C. Figure 8 presents the sentinel sites locations for all sentinel populations.

Figure 8. Sentinel sites and sentinel populations – unlinked anonymous HIV prevalence monitoring sentinel surveillance programme, Slovenia, 2002-2011



LEGEND: Health regions: CE-Celje; KP-Koper; KR-Kranj; LJ-Ljubljana; MB-Maribor; MS-Murska Sobota; NG-Nova Gorica; NM-Novo mesto; RA-Ravne; IDU- injecting drug users; MSM-men who have sex with men; STI-sexually transmitted infections.

At the end of each sampling period, after each sampling year, all serum specimens are tested in pools of 12 for the presence of anti HIV-1/0/2 antibodies using third generation enzyme immuno assay (EIA). Individual sera from reactive pools are re-tested using the same assay. Saliva specimens are tested individually for the presence of anti HIV-1/2 antibodies, using EIA. All EIA repeatedly reactive individual specimens are supplementary tested using HIV Western Blot or Immuno Blot for anti HIV-1 or anti HIV-2 antibodies.

Results

The results for the period of last ten years (2002-2011) are presented in Table 1. MSM are clearly the most affected sub-population.

The proportion of HIV infected is also high among male STI patient, the group that contains a disproportionally high proportion of MSM. The rapid spread of HIV infection among IDU in Slovenia has not started yet. The prevalence of HIV infection among pregnant women, a population at a very low behavioural risk, remains very low. The results for the period 1993-2002 have been published previously (27).

This component of the Slovenian national HIV surveillance system is relatively modest in terms of numbers of tested specimens for all higher-risk behaviour sub-populations, but, the results are informative and provide for crude monitoring of trends in HIV prevalence and early warning.

Table 1. HIV prevalence among sentinel populations of injecting drug users, men who have sex with men, patients with sexually transmitted infections and pregnancies, Slovenia, 2002-2011

	Year	Number of sentinel sites	Number of tested		Number of HIV infected		Proportion of HIV infected	
			Men	Women	Men	Women	Men	Women
Injecting drug users	2002	2	141	41	0	0	0%	0%
	2003	2	253	79	0	0	0%	0%
	2004	3	173	59	0	0	0%	0%
	2005	3	137	57	0	0	0%	0%
	2006	3	125	35	0	0	0%	0%
	2007	3	130	44	0	0	0%	0%
	2008	3	142	34	0	0	0%	0%
	2009	3	127	32	0	0	0%	0%
	2010	4	179	74	1	0	0.6%	0%
	2011	4	136	50	1	0	0.7%	0%
Men who have sex with men	2002	1	113	-	0	-	0.0%	-
	2003	1	101	-	1	-	0.9%	-
	2004	1	79	-	2	-	2.5%	-
	2005	1	82	-	3	-	3.7%	-
	2006	1	94	-	2	-	2.1%	-
	2007	1	124	-	3	-	2.4%	-
	2008	1	137	-	3	-	2.2%	-
	2009	1	117	-	1	-	0.9%	-
	2010	1	114	-	3	-	2.6%	-
	2011	1	105	-	8	-	7.6%	-
Patients with sexually transmitted infections	2002	7	334	201	1	1	0.3%	0.5%
	2003	7	267	200	1	0	0.4%	0%
	2004	7	328	148	5	0	1.5%	0%
	2005	7	403	170	1	1	0.2%	0.6%
	2006	7	419	211	10	0	2.4%	0%
	2007	7	484	257	11	0	2.3%	0%
	2008	7	667	264	23	2	3.4%	0.8%
	2009	6	422	185	13	0	3.1%	0%
	2010	7	525	199	9	0	1.7%	0%
	2011	7	434	198	9	0	2.1%	0%
Pregnancies	2003	8		7544		0		0%
	2005	8		8008		1		0.01%
	2007	8		8963		0		0%
	2009	6		8072		1		0.01%
	2011	7		7231		2		0.03%

When inferring about the distribution and spread of HIV infection in different population groups in Slovenia we should be cautious, as these easily accessible sentinel groups are not representative of all IDU, MSM, patients with STI, and women of reproductive age.

Collating information about the results of testing all donated blood units

All donated blood and blood components have been tested for blood and blood products safety reasons since 1986. Blood donors are a sub-population at a very low risk for HIV infection. We regularly collate information about the results of this mandatory testing for HIV. The proportion of HIV infected donated blood units has remained at a level of approximately one positive result per 100,000 donations. This indirectly indicates a very low level of HIV infection in the general population of Slovenia.

Behavioural surveillance in high-risk groups

In Slovenia, we have managed to develop a very basic behavioural surveillance system in two groups at high behavioural risk for HIV infection, MSM and IDU. The methods and some results are presented for the most affected sub-population in Slovenia, MSM.

Methods

In 2000, we have attached the behavioural data collection to the HIV prevalence monitoring with unlinked anonymous testing among MSM described previously (27). These are small scale, annually repeated one-day cross-sectional surveys in a community setting, just one sentinel site – MSM venue in Ljubljana. MSM themselves organize consecutive sampling of MSM attending an event. In addition to saliva specimens collection for HIV unlinked anonymous testing all participants are invited to anonymously complete a very short self-administered questionnaire. This information is not linked to the saliva specimen of the same individual and the HIV testing result, which cannot be reported to the individual who provided the specimen. Safer sex is promoted and information about access to voluntarily confidential and also anonymous HIV testing and counselling is provided to all MSM included into the sentinel surveillance sample. We collect information on:

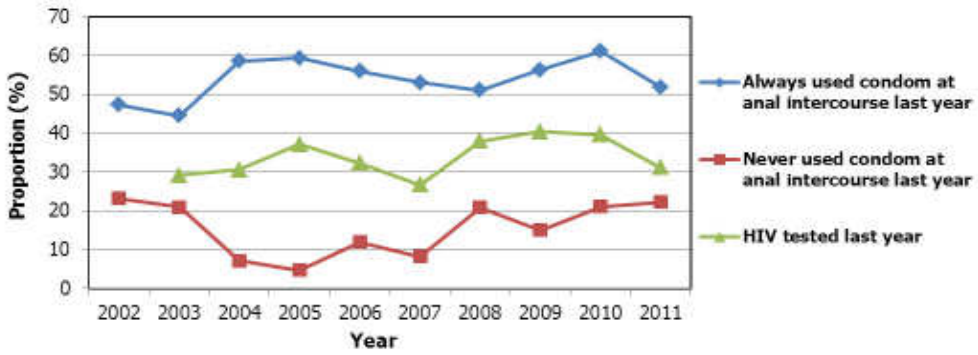
- condom use at anal and oral sex (last occasion and during the preceding year);
- number of anal and oral sex male partners last year;
- number of female partners and condom use with female partners (the bridging to the general population indicator);
- having paid for and having been paid for homosexual sex last year;
- having tested for HIV last year and age in two broad categories (under 25, 25+ years).

Results

Figure 9 presents the variation in proportion of MSM who have reported different frequencies of condom use at anal sex with men during the year preceding the survey.

The proportion of those who reported to have always used condoms has increased from below 50% during 2002-2003 to above 50% during 2004-2011. These results suggest that safer sex behaviour among MSM in Ljubljana has not deteriorated substantially during recent years. Regretfully, the convenient samples sizes have been very small. However, this very simple and crude monitoring of behaviour change among MSM provides informative results with very modest resources.

Figure 9. Condom use at anal sex of men with men during last year, Ljubljana, 2002-2011



Information complementing surveillance information

Surveillance information is often triangulated or complemented with information from ad hock epidemiological studies.

To formulate appropriate and effective sexual health and reproductive health policies, including prevention of HIV, it is crucial to understand sexual behaviour of the population. An example of such a study in Slovenia was the first national sexual behaviour survey conducted in a probability sample of the general population.

Two broadly defined objectives were:

- to describe the patterns of sexual behaviour and identify demographic, social, and behavioural determinants of higher risk behaviour patterns, and

- to describe the distribution of sexually transmitted Chlamydia trachomatis infection.

Several results have been published (28-31). Information about condom use at first heterosexual intercourse is of particularly great relevance to HIV prevention (32). The information presented was similar to surveillance results, as it reflected changes through time (32).

Methods

Methods' details were published previously (33). In brief, we included Slovenian citizens 18-49 years old. We used stratified two-stage sampling. The sampling frame was designed using the list of enumeration areas and information from the Central Population registry.

The data was collected in 1999-2001 by means of face-to-face interviews in combination with anonymous self administration of more sensitive questions in the presence of interviewers (pencil and paper). We adapted the data collection methods used in the first National Sexual Attitudes and Lifestyles survey conducted in 1990 and 1991 in Britain (34). At the end of the interview each respondent was invited to provide first void urine specimen to be confidentially tested for sexually transmitted Chlamydia trachomatis infection.

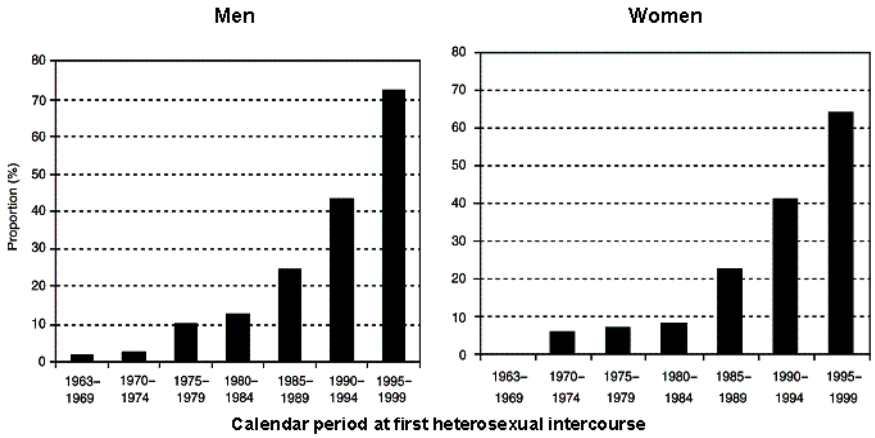
Analyses were conducted using the methods for complex survey data in statistical package STATA. All estimates were weighted.

Results

Figure 11 presents one result which is especially relevant for HIV prevention. Separately for men and women, the variation in the proportion of respondents who reported condom use at first heterosexual intercourse according to the time period of the event is shown. Among those who experienced first heterosexual intercourse during 1970-1974, only 2% of men and 6% of women reported to have used a condom. In contrast, among those who experienced first heterosexual intercourse during 1995-1999, already 72% of men and 64% of women reported condom use. The increase in condom use was most marked during the late 1980s and 1990s, the period during which the effect of "exposure" to AIDS awareness and condom use promotion was increasing.

In Slovenia, we have not missed the window of opportunity for safer sex promotion including condom use promotion targeted to the general population and youth.

Figure 11. Proportion of men and women reporting condom use at first heterosexual intercourse according to calendar period the sexual debut occurred, Slovenia



Strengths and limitations of the Slovenian HIV surveillance system

The Slovenian HIV surveillance system generates information that allows a fairly good understanding of the national burden of HIV, the identification of sub-populations most affected and most at risk as well as monitoring trends in the evolving HIV epidemic and underlying risk behaviours. Based on HIV surveillance information priorities for public health interventions have been identified and resources allocated correspondingly. Surveillance information has provided insight into the impact of public health interventions.

The HIV surveillance system is fairly simple, provides information of fair quality, and seems to be acceptable to participating persons and organizations. The data collection methods have been relatively stable over time, enabling consistency in monitoring trends. HIV surveillance reports are published annually, which is timely enough (18). In addition, short quarterly reports are also published to provide for early warning in case of drastic increases in the numbers of newly diagnosed cases of HIV. The resources used to coordinate the surveillance system used at the Institute of Public Health of the Republic of Slovenia have been very modest.

The information on reported AIDS cases is relatively complete. AIDS is a serious disease, the individual presumably seeks health care and is very likely to be diagnosed correctly and when so, the case is most likely reported. In contrast, HIV infection is asymptomatic for many years and can often be diagnosed very late. Thus, the information on newly diagnosed HIV case most often can not reflect the true incidence of HIV infection.

The information obtained from monitoring HIV prevalence with unlinked anonymous testing in sentinel groups at different behavioural risk (the sentinel surveillance component) has been informative and provided for crude monitoring of trends and early warning. The strengths of such sentinel surveillance include: feasibility, consistency, minimal participation bias, anonymity, and no need for additional invasive procedures to obtain biologic specimens. The limitations include: non-representativeness for all IDU, MSM, patients with STI and pregnant women (low geographical coverage and selection bias), too small sample sizes among MSM and IDU to be able to reliably monitor smaller changes in HIV prevalence, and non-availability of additional risk information (e.g. it is not known whether a MSM or a patient with and STI also had a history of IDU).

The information obtained from the very small scale and simple behavioural surveillance among MSM provides informative results with very modest resources. Since preventing the spread of HIV infection depends mostly on preventing high-risk behaviour and supporting behavioural change, monitoring behaviour is a necessary component of any HIV surveillance system and provides the information for evidence-based targeting of prevention interventions and monitoring their impact. The validity of self-reported information can always be questioned, but, if it does not change with time, such approach is good for monitoring trends. The strengths of such sentinel behaviour surveillance include: feasibility, consistency, anonymity that may contribute to the validity of self-reported information. The limitations include: non-representativeness for all MSM (low geographical coverage and selection bias), participation bias, too small sample sizes to be able to reliable monitor smaller changes in behavioural patterns, and limited amount of behavioural information collected.

Main challenges to improve the Slovenian HIV surveillance system

The existing HIV surveillance system fails to answer many relevant questions. Some examples of such questions are:

- How to prevent the spread of HIV among MSM effectively?
- Are there smaller, but important changes in the high-risk behaviour patterns among MSM?
- What is the proportion of undiagnosed HIV infection in the whole population and in the identified groups at higher behavioural risk?
- What is the proportion of diagnosed HIV infections not reported in the whole population and in the identified groups at higher behavioural risk?
- What is the uptake of HIV testing in different population groups at higher behavioural risk?
- What is the level of HIV testing among patients with diseases indicating high-risk behaviours or diseases indicating HIV infection?
- Do smaller changes in HIV prevalence occur that can not be detected by the behavioural surveillance because of relatively small numbers of participants from high-risk groups?
- Do changes in HIV prevalence occur among MSM and IDU in parts of the country that are not covered with the sentinel HIV surveillance system?
- How many individuals with HIV infection die from AIDS and how many from other causes of death and from which other causes?
- How many individuals with HIV infection have access to highly-active antiretroviral therapy?
- How many individuals with HIV infection are infected with HIV strains resistant to certain anti-retroviral drugs?

In addition to sustaining the existing Slovenian second generation HIV surveillance system challenges include:

- information collected through HIV and AIDS case reporting should include information on HIV therapy and resistance to anti-retroviral drugs;
- information collected through reports of deaths among HIV infected should include more detailed information on causes of death;
- the timeliness and completeness of HIV case reporting should be improved by the introduction of mandatory reporting of diagnoses of HIV infection from laboratories;
- laboratory based HIV surveillance should provide information on the uptake of HIV testing in different population groups at higher behavioural risk and on the level of HIV testing among patients with diseases indicating high-risk behaviour or indicating HIV infection. HIV prevalence monitoring in sentinel high-risk behavioural groups should be improved by increasing the geographical coverage and increasing sample sizes to provide for the detection of smaller changes in HIV prevalence;

- behavioural surveillance among MSM should be improved by regularly repeating large scale in-depth behavioural surveys with integrated biological markers (e.g. other STI).

Finally, formal evaluation of the Slovenian HIV surveillance system should be conducted using the guidelines published by the Centers for Disease Control and Prevention (2). The evaluation findings should be used to improve the efficiency and effectiveness of the surveillance system.

Conclusions and recommendations for prevention, treatment and care

Slovenia is a low HIV epidemic country with les than one individual living with HIV infection per 1000 population. MSM are the most affected sub-population. Rapid spread of HIV infection has not started yet among IDU and their sexual partners. We have not missed the window of opportunity for safer sex promotion including condom use promotion targeted to the general population and youth. We have also managed to target HIV prevention at groups with the highest behavioural risk, particularly MSM.

Prevention and control of HIV infection within the broader frame of promoting sexual and reproductive health remains an important public health priority.

As MSM are the most affected population in Slovenia, promotion of responsible and safer sex including condom use is especially important in this group and should be implemented by MSM non-governmental organizations. Sufficient resources for good coverage of all MSM with preventive interventions should be available and the interventions should be monitored for quality and impact.

In addition, prevention and harm reduction interventions should be targeted to other population groups at higher behavioural risk such as IDU, sex workers and their clients, prisoners, travellers to countries with generalised epidemics and immigrants from these countries to Slovenia.

Voluntary confidential testing for HIV should be promoted in population groups at higher behavioural risk, especially MSM.

Health-care provider initiated testing should be routinely offered to all patients with conditions indicating high risk behaviour (e.g. STI, hepatitis C) or indicating HIV infection (e.g. tuberculosis, one of the AIDS indicator diseases) (35).

Health care case management of all STIs should include counselling for safer sex, notification and treatment of contacts and routine offer for voluntary confidential testing for HIV.

To contain the mortality and morbidity among people living with HIV on a low level, universal access to high quality treatment including highly-active anti-retroviral therapy should sustained.

Exercises

Task 1

Students read the theoretical background of the module, the case study, and recommended readings.

Task 2

Students search for other published surveillance results from their own country (e.g. in Slovenia: Biomedicina Slovenica; COBISS-Cooperative Online Bibliographic System of Slovenia; Institute of Public Health of the Republic of Slovenia) and from other European countries (e.g. on web-sites of other national public health institutes such as Health Protection Agency in the UK; Eurosurveillance website; and from international bibliographic data-bases such as Medline and PubMed) and reads a few examples.

Task 3

Each student presents an example of a national surveillance system and the whole group discusses the objectives of the presented surveillance systems, methods used, results and uses of surveillance information, characteristics of the surveillance system, its strengths and limitations as well as proposals for improvements.

References

1. Last JM. A dictionary of Epidemiology. Oxford: Oxford University Press, 2001.
2. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50 (No. RR-13):1-36. Available at: URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm (Accessed 19 July 2008).
3. Berkelman RL, Stroup DF, Buehler JW. In: Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press, 2002. p 759-78.
4. Institute of Medicine. The future of public health. Washington: National Academy Press, 1988.
5. Hadden F, O'Brien S. Assessing acute health trends: surveillance. In: Pencheon D, Guest C, Meltzer D, Gray JAM (editors). Oxford handbook of public health practice. Oxford: Oxford University Press, 2005. p 14-9.

6. Grgic-Vitek M, Klavs I, Potocnik M, Rogl-Butina M. Syphilis epidemic in Slovenia influenced by syphilis epidemic in the Russian Federation and other newly independent states. Int J AIDS & STD 2002;13 (Suppl 2):2-4.
7. Grgic-Vitek M, Klavs I, Kraigher A. Re-emergence of pertusis in Slovenia: Time to change immunization policy. Vaccine 2008;26;1874-8.
8. Institute of Public Health of the Republic of Slovenia. Health statistics annual report 2008. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia. Available at: URL: http://www.ivz.si/Mp.aspx?ni=78&pi=6&_6_id=407&_6_PageIndex=0&_6_groupId=-2&_6_newsCategory=IVZ+kategorija&_6_action=ShowNewsFull&pl=78-6.0. (Accessed 22 March 2013)
9. Institute of Public Health of the Republic of Slovenia. Out-of hospital care statistics. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia. Available at: URL: http://www.ivz.si/Mp.aspx?ni=78&pi=6&_6_id=65&_6_PageIndex=0&_6_groupId=-2&_6_newsCategory=IVZ+kategorija&_6_action=ShowNewsFull&pl=78-6.0. (Accessed 22 March 2013)
10. Institute of Public Health of the Republic of Slovenia. Health-care information system on hospitalizations. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia. Available at: URL:

    http://www.ivz.si/Mp.aspx?ni=78&pi=6&_6_id=2141&_6_PageIndex=0&_6_groupId=-2&_6_newsCategory=IVZ+kategorija&_6_action=ShowNewsFull&pl=78-6.0. (Accessed 22 March 2013)
11. Institute of Public Health of the Republic of Slovenia. Perinatal information system and surveillance system for perinatal deaths. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia. Available at: URL: http://www.ivz.si/Mp.aspx?ni=78&pi=6&_6_id=915&_6_PageIndex=0&_6_groupId=-2&_6_newsCategory=IVZ+kategorija&_6_action=ShowNewsFull&pl=78-6.0. (Accessed 22 March 2013)
12. Institute of Oncology Ljubljana. Cancer Registry of Slovenia. (in Slovene) Ljubljana: Institute of Oncology Ljubljana. Available at: URL: http://www.onkoi.si/eng/dejavnosti/epidemiologija_in_register_raka/registri_raka/register_raka_rs/ (Accessed 22 March 2013)
13. Institute of Oncology Ljubljana. Cancer in Slovenia 2009. Ljubljana: Institute of Oncology Ljubljana, 2013. Available at: URL: http://www.onko-i.si/fileadmin/onko/datoteke/dokumenti/RRS/LP_2009.pdf (Accessed 22 March 2013)
14. Slovenian Parliament. Communicable Diseases Law. (in Slovene) Official Journal of the Republic of Slovenia No 69/1995.
15. Ministry of Health. Regulation on notification of communicable diseases and special activities for prevention and control. Official Journal of the Republic of Slovenia No 16/1999.
16. Slovenian Parliament. Law on Health Care Data Bases. (in Slovene) Official Journal of the Republic of Slovenia No 65/2000.
17. Institute of Public Health of the Republic of Slovenia. Surveillance of communicable diseases in Slovenia, 2011. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia. 2012. Available at: URL:

    http://www.ivz.si/gradiva_nalezljive_bolezni?pi=5&_5_Filename=6179.pdf&_5_MediaId=6179&_5_AutoResize=false&pl=105-5.3. (Accessed 22 March 2013)
18. Klavs I, Kustec T, Kastelic Z. HIV infection in Slovenia: annual report 2011. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia, 2012. Available at: URL: http://www.ivz.si/hiv_spo?pi=5&_5_Filename=5000.pdf&_5_MediaId=5000&_5_AutoResize=false&pl=107-5.3. (Accessed 22 March 2013)
19. Klavs I, Kustec T, Kastelic Z. Sexually transmitted infections in Slovenia: annual report 2011. Ljubljana: Institute of Public Health of the Republic of Slovenia, 2012. Available at: URL: http://www.ivz.si/hiv_spo?pi=5&_5_Filename=5502.pdf&_5_MediaId=5502&_5_AutoResize=false&pl=107-5.3. (Accessed 22 March 2013)
20. Ucakar V, , Sevljak-Jurjevec M, Grgic-Vitek M, Javornik S, Kraigher A. Analysis of the implementation of the immunisation programme in Slovenia in 2010. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia, 2011. Available at: URL: http://img.ivz.si/janez/357-4943.pdf (Accessed 22 March 2013)

21. Ucakar V, Jeraj I, Grgic-Vitek M, Sevljak-Jurjevec M, Kraigher A. Adverse vaccination events in Slovenia in 2011. (in Slovene) Ljubljana: Institute of Public Health of the Republic of Slovenia, 2012. Available at: URL: http://img.ivz.si/janez/378-5676.pdf (Accessed 19 October 2013).
22. World Health Organization (WHO) and Joint United Nations Programme on HIV/AIDS (UNAIDS). Guidelines for the second generation HIV surveillance. Geneva: WHO and UNAIDS, 2000. Available at: URL: http://www.who.int/hiv/pub/surveillance/pub3/en/ (Accessed 19 October 2013)
23. WHO and UNAIDS. Initiating second generation HIV surveillance systems: practical guidelines. Geneva: WHO and UNAIDS, 2002. Available at: URL:

    http://www.who.int/hiv/pub/surveillance/guidelines/en/ (Accessed 19 October 2013)
24. WHO. Guidelines for effective use of data from HIV surveillance systems. Geneva: WHO and UNAIDS, 2004. Available at: URL:
    http://www.who.int/hiv/strategic/surveillance/hivpubsurveillance/en/ (Accessed 19 October 2013)
25. European Centre for the Epidemiological Monitoring of AIDS. 1993 revision of the European AIDS surveillance case definition. AIDS surveillance in Europe. Paris: institute de Medicine et D'Epidemiology Africaines. 1993 Quarterly Report No.37.
26. EuroHIV. HIV/AIDS Surveillance in Europe. End-year report 2006. Saint-Maurice: Institute de Veille Sanitaire. 2007;75.
27. Klavs I, Poljak M. Unlinked anonymous monitoring of HIV prevalence in high and low-risk groups in Slovenia, 1993-2003. Croat Med J 2003;44:545-9.
28. Klavs I, Rodrigues LC, Wellings K, Kese D, Hayes R. Prevalence of genital Chlamydia trachomatis infection in the general population of Slovenia: serious gaps in control. Sex Transm Infect 2004;80:121-3.
29. Grgic-Vitek M, SvaB I, Klavs I. Prevalence of and risk factors for self-reported sexually transmitted infections in Slovenia in 2000. Croat Med J 2006;47:722-9. Available at: URL:
    http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2080463 (Accessed 19 October 2013)
30. Klavs I, Rodrigues LC, Weiss HA, Hayes R. Factors associated with early sexual debut in Slovenia: results of a general population survey. Sex Transm Infect 2006;82:478-83.
31. Klavs I, Hamers FF. Male circumcision in Slovenia: results from a national probability sample survey. Sex Transm Infect 2008;84:49-50.
32. Klavs I, Rodrigues LC, Wellings K, Weiss HA, Hayes R. Increased condom use at sexual debut in the general population of Slovenia and association with subsequent condom use. AIDS 2005;19:1215-23.
33. Klavs I, Kese D, Svab I. Slovene national survey of sexual lifestyles, attitudes and health, 1999-2001: data collection methods. Zdrav Vars 2007;46:1-8.
34. Johnson AM, Wadsworth J, Wellings K, Field J (editors). Sexual Attitudes and Lifestyles. Oxford: Blackwell Scientific publications, 1994.
35. WHO. Guidance on provide-initiated HIV testing and counselling in health care facilities. Geneva: WHO, 2007.

Recommended readings

1. Berkelman RL, Stroup DF, Buehler JW. In: Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press, 2002. p 759-78.
2. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50(No.RR-13):1-36. Available at: URL:
    http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm (Accessed 19 October 2013).
3. Hadden F, O'Brien S. Assessing acute health trends: surveillance. In: Pencheon D, Guest C, Meltzer D, Gray JAM (editors). Oxford handbook of public health practice. Oxford: Oxford University Press, 2005. p 14-9.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Delphi analysis**
Module: 2.19	ECTS (suggested): 0.25
Author(s), degrees, institution(s)	Neda Milevska-Kostova, MSc, MCPPM, Executive Director Centre for Regional Policy Research and Cooperation "Studiorum", Skopje, Macedonia William N. Dunn, PhD, Professor of Public and Urban Affairs, Graduate School of Public and International Affairs (GSPIA), University of Pittsburgh, PA, USA
Address for Correspondence	Neda Milevska-Kostova, MSc, MCPPM, Executive Director Centre for Regional Policy Research and Cooperation "Studiorum" Nikola Parapunov 41, kompleks Makoteks 2 kat, Skopje, Macedonia e-mail: nmilevska@studiorum.org.mk
Keywords	Delphi study, Delphi technique.
Learning objectives	After completing this module students and public health professionals should be able to: • distinguish projections and predictions; • contrast potential, plausible, and normative futures; • evaluate judgmental forecasting methods; • use policy Delphi analysis to make point and interval forecasts; • create a Delphi forecast for an issue.
Abstract	This module focuses on Policy Delphi analysis, which is one of the most important and widely used methods of employing expert judgment to make forecasts. The module gives overview to the three forms of forecasting used for Delphi analysis, and explores the techniques used in these types of forecasts. Based on the described assessment of the strengths and limitations of policy Delphi analysis, the module also describes the outline of the use of a 2-round Delphi exercise in class. The purpose is to forecast the probable outcomes in the pre-defined health policy case. In this module, the teachers are offered a number of sample cases that they can use or adapt for use as class exercises.
Teaching methods	An introductory lecture gives the students insight into characteristics of the Delphi technique. The theoretical knowledge is illustrated with a pre-developed case study. Students are required to read some recommended readings before the introductory lectures. After the introductory lectures, they are given the assignment to prepare a Delphi forecast design, based on the theoretical background, which they discuss with other students, especially the designing and planning phase and the expected outcomes of the Delphi analysis. In continuation, they need to create a virtual Delphi analysis, using published materials on the selected health issue and from their own knowledge of the problem, and present their findings to other students in the class.
Specific recommendations for teachers	• ECTS: 0.25; • work under teacher supervision/individual students' work proportion: 40%/60% ; • facilities: lecture room; • equipment: LCD projection, whiteboard (flipchart); • training materials: recommended readings; • target audience: master degree students according to the Bologna process.
Assessment of students	Multiple-choice questionnaire (MCQ) and case problem presentations.

DELPHI ANALYSIS

Neda Milevska-Kostova, William N. Dunn

Theoretical background

History and objectives of the Delphi method

A relatively old but still fairly often used method for objective and reliable exploration of new ideas through a consultation of independent experts in a creative way of forecasting to furnish arguments for decision making - is a short compiled definition of the Delphi method, which was developed in the 1940s by the RAND Corporation, as a tool that evolved in the process of experimentation in the technology forecasting studies. In 1944, General Arnold asked Theodor von Karman to prepare a forecast of future technological capabilities that might be of interest to the military (1). Later, in 1959 Helmer and fellow RAND researcher Rescher published a paper on "The Epistemiology of the Inexact Sciences", which provided the philosophical base for forecasting (2). The paper argued that in fields in which the instruments are not yet developed to the point of scientific laws, the testimony of experts is permissible and should be acceptable. Thus the Delphi method was developed, recognizing human judgement as legitimate and useful input in generating forecasts - solving the problem of how to use human judgement and, specifically, how to combine the testimony of a number of experts into a single useful statement (3).

Objectives of the Delphi method

The Delphi method is a judgemental forecasting technique for obtaining, exchanging and developing informed opinion through a consensus about the most probable future by iteration.

Characteristics of the Delphi method

The Delphi method is an exercise in-group communication among a panel of geographically dispersed experts (4). The technique allows experts to systematically deal with a complex problem or task. In its essence, the Delphi technique is rather straightforward, with set rules and principles of its performance; it comprises sets/series of questionnaires sent to individuals and independent experts, which have been pre-selected for the task. According to Fowles (2) anonymity, controlled feedback, and statistical response characterize Delphi. The group interaction in Delphi is anonymous, in the sense that comments and forecasts are not linked to their generator, but are presented to the group in such a way as to suppress any identification (3).

Principles of the Delphi method

Anonymity

Anonymity is the first and by far the most important principle of the Delphi. This approach allows for equality among the participating experts, avoiding the positions of authority or dominance of one's opinion or judgement. The panel of experts is selected based on their knowledge or opinion on the issue.

The questionnaires are designed is such a way to provoke and develop individual responses to the problems posed and to enable the experts to refine their views within the progress of the group work towards the goal of the task. The main point behind the Delphi method is to overcome the disadvantages of conventional committee action (3).

Iteration

Another important principle of the Delphi is the possibility of iteration - repetitiveness of the process as long as the experts, and especially the group coordinator, feel that there is a need for refinement of the views and statements. This principle, however, does not in any way imply that all experts necessarily have to agree on certain point (see the principle of stakeholders disagreement below), but reaching a point at which all participants feel comfortable with the outcome is a necessary precondition for considering the Delphi analysis to have undergone is a successful manner.

The rounds of exchanging judgements can be repeated as required, but most often the process is completed after two or three cycles; with either the expert consensus (see below), which is common for the traditional Delphi approach, or with stakeholder disagreement (see below), which is a newer principle often related to the policy Delphi approach.

Controlled feedback

The process is coordinated by a person called facilitator, who needs to have both analytical and managerial skills, as s/he is responsible for sending out questionnaires with instructions, collecting them back and summarizing views, as well as preparing a new set of questions that would serve as a further distiller of the ideas and opinions, if and when consensus is not reached during the first or subsequent

rounds. His/her analytical skills are needed for the process of synthesis of results and working towards building a group consensus.

Statistical group response

In order to represent the full range of opinions and not only the ones reached by consensual agreement of the experts panel, the summary of individual responses are presented in a form of measures of central tendency (usually the median), dispersion (the interquartile range) and frequency distributions (histograms and frequency polygons) (4,5).

Expert consensus

Traditionally, the Delphi method has aimed at a consensus on the most probable future by iteration; the number of cycles to reach the consensus was irrelevant, as long as there is obvious progress in moving towards the anticipated consensus. However, the weakness of this aspect is that not always the consensus can be reached, regardless of the number of repetitions, leading to inevitable changes in the selected expert panel; this imperfection of the method, can potentially weaken the interest of the remaining expert members, as a result of the appearing sense of time being wasted. Thus, in the late 1960s alternative Delphi approaches were developed and introduced.

Stakeholders' disagreement

The Policy Delphi (6) launched by Murray Turoff instead is a decision support method aiming at structuring and discussing the diverse views of the preferred future; the Policy Delphi, seeks to generate the strongest possible opposing views on the potential resolutions of a major policy issue. In the author's view, a policy issue is one for which there are no experts, only informed advocates and referees (6).

Policy Delphi begins by using snowball sampling to maximize (rather than minimize) differences among multiple perspectives, using these differences to inform the development of consensus and predictive accuracy in policy forecasting (5,7-9). In the face of the policy issue, the expert becomes an advocate for effectiveness or efficiency and must compete with the advocates for concerned interest groups within the society or organization involved with the issue. The Policy Delphi also rests on the premise that the decision maker is not interested in having a group generate her/his decision; but rather, have an informed group present all the options and supporting evidence for her/his consideration (9).

Other types of Delphi

Besides the traditional and the policy Delphi, several other modified types have been proposed.

The Argument Delphi (10) developed by Osmo Kuusi focuses on the ongoing discussion and finding relevant arguments rather than on the output of the debate itself. The process is based on a four-level classification of statements, which are usually at least in part mutually exclusive; experts at first make a very simple evaluation on whether they approve or disapprove the statement/topic, after what they develop arguments pro et contra for acceptance or rejection of a given statement.

The Disaggregative Policy Delphi (11) developed by Petri Tapio uses cluster analysis as a systematic tool to construct various scenarios of the future in the latest Delphi round. The respondent's view on the probable and the preferable future are dealt with as separate cases.

Forecasting used in health policy

There are several types of forecasting techniques used in the health policy development; they often employ judgemental or statistical methods of forecasting. The statistical methods (extrapolation, multivariate forecasting and econometric forecasting) are used when there is sufficient statistical data; however, when there are insufficient data sources or there is low emphasis on the accuracy of the forecast, experts use judgemental methods, such as the unstructured (unaided) or structured, of which Delphi is one example.

The Delphi technique is used often in the health sector when there is insufficient or unreliable data to conduct a statistical forecast. Projections developed by Delphi panels are believed to be more accurate than forecasts based on unaided judgment. There is limited direct evidence of the accuracy of forecasts using the Delphi method (12).

Steps of Delphi

Conventional (traditional) Delphi

A Delphi method, as defined by its creators (2) has the following 10 steps:

1. Formation of a Delphi team to undertake and to monitor the project.
2. Selection of one or more panels to participate in the exercise. Customarily, the participants are experts in the investigation area.
3. Development of the first round Delphi questionnaire.
4. Testing the questionnaire for proper wording (e.g., ambiguities, vagueness).
5. Transmission of the first questionnaires to the panellists.

6. Analysis of the first round responses.
7. Preparation of the second round questionnaires (and possible testing).
8. Transmission of the second round questionnaires to the panellists.
9. Analysis of the second round responses. (Steps 7 to 9 are reiterated as long as desired or necessary to achieve stability in the results.)
10. Preparation of a report by the analysis team to present the conclusions of the exercise.

However, both the creators and other authors (13) argue that the most important issue in this process is the understanding of the aim of the Delphi exercise by all participants. Otherwise, the panellists may answer inappropriately or become frustrated and lose interest (3).

On the other hand, the Policy Delphi can be conducted in a number of different ways, depending on the context and the skill and ingenuity of the persons using the technique. Since Policy Delphi is a major research undertaking, it involves a large number of technical questions, sampling, questionnaire design, reliability and validity, and data analysis and interpretation (5). The steps of the Policy Delphi, which are somewhat modified from the conventional approach, are:

1. issue specification,
2. selection of advocates,
3. questionnaire design,
4. analysis of first-round results,
5. development of subsequent questionnaires,
6. organization of group meetings,
7. preparation of final report.

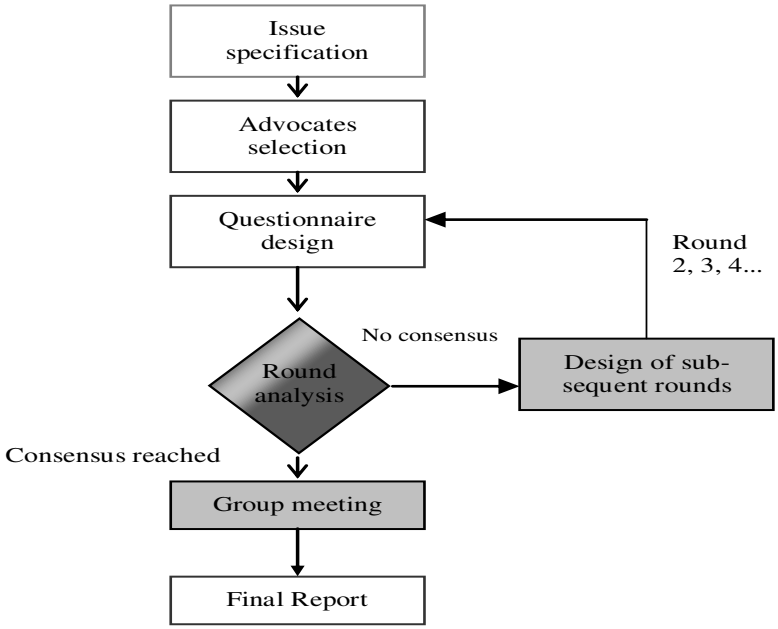
As the Policy Delphi is more often used for health policy development, its steps are elaborated in the Table 1.

Table 1. Steps of Policy Delphi (6)

STEP	DESCRIPTION	EXAMPLE*
Issue specification	Analyst decides upon the specific issues for which the Delphi will be conducted.	To develop a range of possible national drug-abuse policy options.
Selection of advocates	Selection of key stakeholders in the issue area, including experts (preferably with opposing or conflicting positions on the issue).	A list of about 100 experts was set, with invitations initially sent to 45 persons, for an expected positive response from 25 persons (in fact, positive responses obtained from 38 experts).
Questionnaire design	Development of questionnaire based on the selected issue; depending on the level of knowledge of the analyst and selected experts, the questionnaire is either done completely by the analyst (using scales for measurement, see example below) or the analyst develops an open-end questions set, that are further distilled by the experts through providing their opinions.	The first questionnaire was pretested, and the entire forecasting section deleted when it was determined that the time to complete the questionnaire was decreased considerably by deleting this section. The questionnaire consisted of 4 sections: development of objectives, transition matrix, policy issue statements and additional items.
Analysis of first-round results	Analyst attempts to determine the initial positions on the forecasts, issues, goals and opinions. Use of summary measures due to expected conflicting assessments of various advocates (avoiding presentation of central tendency only).	24 of 35 respondents (69%) returned the filled-in questionnaire; respondents were asked substantive-issue questions, self-rating of own expertise in drug abuse, expectations from the study, etc. Experts listed 78 policy issues that were distilled down to 55; the 187 key indicators were culled to 153.
Development of sub-sequent	The results of prior rounds are used for development of	The 2nd questionnaire developed after round One was completed; it

STEP	DESCRIPTION	EXAMPLE*
questionnaires	subsequent ones; include summary of arguments for the most conflicting judgements. The analyst should count on decreased rate of response in the subsequent rounds, especially those involved on voluntary basis.	included only two sections: National drug-abuse policy objectives and Policy issue statements. There were four issues that exhibited marked differences between policy experts and nonexperts in the importance of issues.
Organization of group meetings	Bringing advocates (stakeholders) in face-to-face discussion of reasons, assumptions and arguments of various positions; useful for immediate feedback.	-
Preparation of final report	Analyst is responsible for drafting the final report, based on the questionnaire results and face-to-face discussions; the final report reveals a review of various issues and options available, including a complete description of all conflicting positions.	Final report covering the main conclusions of the questionnaire analysis, as the group meting was not held in this example.

Figure 1. Policy Delphi steps and process



Design of questionnaire

First questionnaire

As explained earlier, the Delphi usually consists of several rounds; of all, only the first questionnaire can be prepared in advance and all other are derived based on the synthesized results of previous rounds.

In the first round, the Delphi process traditionally begins with an open-ended questionnaire. The open-ended questionnaire serves as the cornerstone of soliciting specific information about a content area from the Delphi subjects (14). For example, in the first questionnaire, participants might be asked to provide their judgment (15) on a most probable period by which a drug abuse policy will give visible effects. After receiving subjects' responses, investigators convert the collected information into a well-structured questionnaire. This first-round questionnaire is used as the survey instrument for the second round of data collection. It should be noted that it is both an acceptable and a common modification of

the Delphi process format to use a structured questionnaire in Round 1 that is based upon an extensive review of the literature (16). Kerlinger (17) noted that the use of a modified Delphi process is appropriate if basic information concerning the target issue is available and usable.

Results analysis and providing feedback

Analysis of the results

When questionnaires are collected, the analyst or the facilitator is the one that summarizes the responses, synthesizes various positions and puts forward the arguments already pointed out by the advocates. Each sub-sequent round is structured and prepared based on the results of the previous one; however, the results are not presented to the panel of advocates/experts at each round in a form of report, but rather in a form of a new questionnaire (see above Design of questionnaire).

Data analysis can involve both qualitative and quantitative data; usually, qualitative data is dealt with if conventional Delphi studies, which use open-ended questions to solicit subjects' opinions, are conducted in the initial iteration. Subsequent iterations are to identify and hopefully achieve the desired level of consensus among panellists.

The statistics used in Delphi studies are most commonly measures of central tendency (means, median, and mode) and level of dispersion (standard deviation and inter-quartile range) in order to present information concerning the collective judgments of respondents (18). Generally, the uses of median and mode are favored (16).

Table 2. Types of items and scales used in Policy Delphi Questionnaire (5)

TYPE OF ITEM	ITEM	SCALE
Forecast	According to a projection of researchers at the National Public Health Institute, over 20% of young people age 15-24 are smoking marijuana, and this percentage will be doubled in the coming 10 years. How certain are you that this projection is **reliable**?	[1] Certainly reliable [2] Reliable [3] Risky [4] Unreliable [0] No judgement
Issue	Personal use of marijuana should/should not be legalized. How **important** is this issue relative to others?	[1] Very important [2] Important [3] Slightly important [4] Unimportant [0] No judgement
Goal	One goal of National drug abuse policy is to increase the awareness of difference between drug use (responsible) and drug abuse (irresponsible). How **desirable** is this objective?	[1] Very desirable [2] Desirable [3] Undesirable [4] Very undesirable [0] No judgement
Options	Drug abuse education is reported to contribute towards reduction of potential users. How **feasible** ins this policy option?	[1] Definitely feasible [2] Feasible [3] Possibly feasible [4] Definitely unfeasible [0] No judgement

Providing feedback

As mentioned above, the results analysis from each round serves as feed into the next round of challenging experts'/advocates' positions and opinions on the issue; once the issue is exhausted, there is seem no further argumentative development the facilitator/analyst faces for writing the final report. The final report is aimed primarily at informing of all opposing positions and alternatives presented by the advocates/experts, supported by arguments as presented by the panel. Yet, as this may become a long list of different standpoints that might confuse the policy- or decision-maker, other presentations are used, taking into consideration the possibility to present all opinions, including the outlayers. Some are mentioned below:

- difference graph,
- histogram or bar chart,
- box-and-whisker plot,

- lists of measures of central tendency and dispersion, etc.

Figure 2. Providing feedback - Difference graph (5)

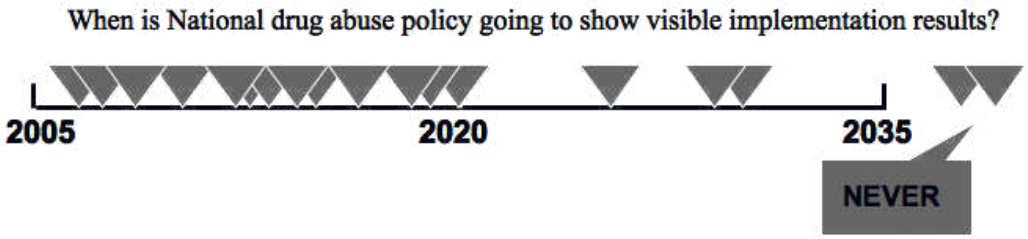
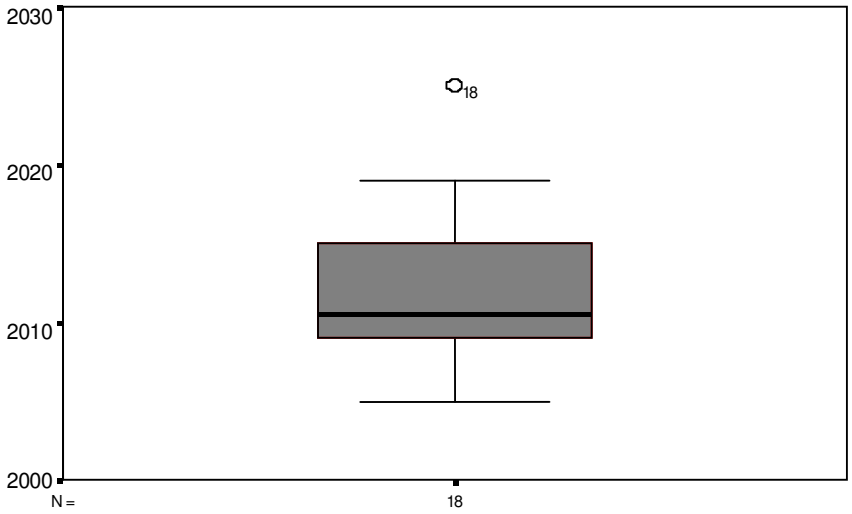


Figure 3. Providing feedback - Whisker plot



On-line tools for Delphi analysis

Available (free) software

The website of Forecasting Principles is offering on-line usable application free of charge for trying up the basic features of the Delphi method. Following the link http://armstrong.wharton.upenn.edu/delphi2/ you can register as administrator, start your own survey, in which the application will help you:

- select experts,
- develop questions and scales,
- obtain responses from the experts,
- summarize a report after each round.

This tool, developed by J. Scott Armstrong (19) is a very useful starter's kit for application of Delphi method into your own research.

Risks and disadvantages of Delphi analysis

Risks associated with Delphi

While offering large possibilities for forecasting by use of experts opinion (Delphi) or advocates positions (Policy Delphi), this method has its own risks, associated mainly with the inseparable personal subjectivism and the partial or complete ignorance of the advocates in the issue. As the outcome of a Delphi sequence is nothing but opinion, and the results of the sequence are only as valid as the opinions of the experts who made up the panel (20), if not carefully selected, the panel of experts, or advocates for that matter, can lead the discussion into prediction of a highly unlikely future outcome; in some cases this can further strengthen the confidence of the part of the experts or advocates which do not possess great knowledge or informed opinion in the issue.

Disadvantages of Delphi

As with other research methods, Delphi has been reported to have its own disadvantages; most of them associated with improper selection or invalid use of the method for a given type of outcome desired. It is up to the researcher/analyst to determine if this method will produce valid and usable results.

The challenges that the researcher should think of prior to selection of this method as their preferred tool, is related to the internal consistency and reliability of judgements among experts, which - if low or lacking - may lead to low reproducibility of forecasts based on the results elicited; sensitivity of results to ambiguity and respondent reactivity in the questionnaires used for data collection; difficulty in assessing the degree of expertise held by participating experts (21).

Another problem identified by researchs into the implementation and application of Delphi surveys has been the tendency for experts to over-simplify particular issues, and treat them as isolated events. This is particularly the case in forecasting, where experts tend to think in terms of linear sequential events, rather than applying a holistic view that involves complex chains and associations. Again, it is up to the researcher/analyst to extract this aspect from the answers of experts/advocates through the sub-sequent rounds of questioning. Other techniques for multi-dimensional analysis are also available, such as 'cross impact matrix forecasting' intended to compare a range of 'possible futures' against each other; but those are discussed elsewhere (4,22,23).

Advantages of Delphi

Although the approach was originally developed to capture expertise in uncertain and emergent domains, it tends to be used in evaluation when significant expertise exists on the subject, for example in the case of programmes that are not innovative. The method is recommended when the questions posed are simple (a programme with few objectives, of a technical nature) and for the purpose of establishing a quantitative estimation of the potential impacts of an isolated intervention (e.g. increase in taxes or in the price of health services). It is also recommended in an ex ante evaluation context if the evaluation concerns public intervention of a technical nature. But, it may also be used to specify relations of causes and potential effects in the case of innovative interventions. It is particularly useful when a very large territory is being dealt with since there are no experts' travel expenses, only communication costs.

It has been found to be particularly useful in programmes related to public health issues (such as, policies for drug use reduction and prevention of HIV/AIDS) and education (1,4). In general, the Delphi method is useful to explore and unpack specific, single-dimension issues. There is less support for its use in complex, multi-dimensional modelling. In these cases, the evidence does suggest that data gathered by Delphi surveys is a useful input, when supported by data gathered from other sources, to complex scenario-building (24).

Nevertheless, according to some authors, still the context plays great role in deciding whether and when to use the Delphi method.

Case study

AIDS vaccination policy: a scenario analysis using the Delphi method[11]

Background

The Delphi method was used to explore and identify the potential implications associated with the introduction of a first AIDS vaccination in Switzerland. Thirty participants with an interest in the field contributed anonymously to the study. The study focused on an existing scenario which modelled the characteristics of a first preventive, partially effective, vaccination against AIDS.

The process

The Delphi consultation was carried out in three stages. In the first round, the participants were asked to:

- list the objectives to be achieved in the first five years;
- evaluate the acceptability and feasibility of proposals concerning the development of a public health strategy and the AIDS vaccination;
- estimate the potential use of the vaccination by different groups of users.

The outcome

The used Delphi method in the consultation process produced two main outcomes: firstly, a set of strategies and recommendations for the development of a framework of AIDS prevention campaigns and, secondly, an institutional framework for the setting up of a future AIDS vaccination strategy.

Follow-up

In parallel, in 2003, clinical trials of a new vaccine against HIV, started in Switzerland and the United Kingdom. EuroVacc, the foundation organising the trials, has tested two vaccines: DNA-C, developed by Professor Hans Wolf of the University of Regensburg, Germany, and its booster, NYVAC,

[11] Source of the case study and data references # 25 and # 26.

developed by the French pharmaceutical company Aventis. About 160 healthy volunteers - half in London and half in the Swiss city of Lausanne, where EuroVacc is based – have been subjected to the test of the vaccine for safety.

In 2005, the combined vaccine consisting of DNA vaccine and NYVAC booster was tested in Switzerland, the Netherlands, the United Kingdom, Spain, Italy, Germany, and Sweden in hundreds of people seen as being at high risk of HIV infection, including gay men, drug users, and commercial sex workers. Volunteers' rate of infection was monitored and compared with the rate of infection in similar groups of people who had not been given the vaccine.

Exercises

Task 1

Before class, based on the Table 2 above, write as many as you can "forecast," "issue," "goal," and "options" questions related to the Case study above. These questions will be collected and used for the in-class Delphi exercise.

Task 2

After the in-class discussion of the collected questions, assume the role given by the teacher and fill out the sample questionnaire that will be distributed by the teacher. One of the students will be assigned the role of analyst who will chair the group meeting in the next class. Role-play exercise in the class.

Task 3

Based on the role-play held in the class, prepare a final report (assuming the role of analyst). The teacher will consider this paper as an assessment for the module.

References

1. Cornish E. The study of the future: an introduction to understanding the art and science for understanding and shaping tomorrow's world. Washington, D.C.: World Future Society; 1977.
2. Fowles J. Handbook of futures research. Greenwood Press: Connecticut; 1978.
3. Illinois Institute of Technology [homepage on the Internet] Chicago: The Delphi method.[cited 2009 Apr 6]. Available from URL:: http://www.iit.edu/~it/delphi.html
4. Adler M, Ziglio E. Gazing into the oracle. Bristol (PA): Jessica Kingsley Publishers;1996.
5. Dunn WN. Public policy analysis. New Jersey: Pearson Prentice Hall; 2004.
6. Turoff M. The design of a policy Delphi. Technological Forecasting and Social Change. 1970;2:2.
7. Mahmoud E. Accuracy in forecasting: a survey. Journal of Forecasting 1984; 3: 139-59.
8. Rescher N. Predicting the future. Albany (NY): State University of New York Press; 1998.
9. Turoff M. The policy Delphi. In: Linstone H, Turoff M, Helmer O. editors. The Delphi method: techniques and applications. Adisson Wesley (NY); 2002. p.80-98.
10. Kuusi O. Expertise in the future use of generic technologies. Research Reports 59. Government Institute for Economic Research Finland (VATT); 1999.
11. Tapio P. Disaggregative policy Delphi: using cluster analysis as a tool for systematic scenario formation. Technological Forecasting & Social Change. 2002;70(1):83-101.
12. Forcastingprinciples.com [homepage on the Internet] Methods for health forecasting, forecasting principles; c1997-2007 [cited 2009 Apr 6]. Available from URL: http://www.forecastingprinciples.com/Health/methods.html
13. Delbecq AL, Van de Ven AH, Gustafson DH. Group techniques for program planning: a guide to nominal group and Delphi processes. Glenview (Ill): Scott, Foresman; 1975.
14. Custer RL, Scarcella JA, Stewart, BR. The modified Delphi technique: a rotational modification. Journal of Vocational and Technical Education. 1999;15(2):1-10.
15. Gordon TJ. The Delphi method. Future research methodology. AC/UNU Millenium Project; 1994.
16. Hsu C-C, Sandford BA. The Delphi technique: making sense of consensus, practical assessment, research and evaluation. 2007;12(10).
17. Kerlinger FN. Foundations of behavioral research. New York: Holt, Rinehart, and Winston, Inc.; 1973.

18. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. Journal of Advanced Nursing. 2000;32(4):1008-15.
19. Imperial College London [homepage on the Internet].The development of the Delphi software was funded in part by International Institute of Forecasters. [cited 2009, Apr 6] Available from URL: http://www3.imperial.ac.uk
20. Martino J. The Precision of Delphi estimates. Technological Forecasting and Social Change. 1970;1(3):293-9.
21. Makridakis S, Wheelwright SC. Forecasting methods for management. 5th ed. Chichester: Wiley; 1989.
22. Gordon TJ, Hayward H. Initial experiments with the cross-impact matrix method of forecasting. Futures. 1968;1(2):100-16.
23. Gatewood RD, Gatewood EJ. The use of expert data in human resource planning: guidelines from strategic forecasting. Human Resource Planning. 1983;5:83-94.
24. European Commission (EC). Evaluating socio economic development. Sourcebook 2: methods & techniques, Delphi method. 2005. p. 110-9.
25. Zuber P. Introducing a first AIDS vaccination in Switzerland: a Delphi policy analysis. Institut de medecine sociale et preventive. Lausanne; 1994.
26. Fleck F. Trials of AIDS vaccine to start in Switzerland and United Kingdom. BMJ 2003;326:952.

Recommended readings

1. Dunn WN. Public policy analysis. New Jersey (NY) Pearson Prentice Hall; 2004 p. 180-98.
2. Turoff M. The Policy Delphi. In: Linstone H, Turoff M, Helmer O. editors. The Delphi method: techniques and applications. New York: Addison-Wesley; 2002 p. 80-98.
3. Jillson IA. The National drug abuse policy Delphi: Progress report and findings to date. In: Linstone HA, Turoff M, Helmer O. editors. The Delphi method: techniques and applications. New York: Addison-Wesley; 2002. p. 124-59
4. Zuber P. Introducing a first AIDS vaccination in Switzerland: a Delphi policy analysis. Institut de medecine sociale et preventive. Lausanne; 1994.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**The Geographic Information System (GIS): The example of traffic air pollution**
Module: 2.20	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Andreja Kukec, Bsc, PhD Candidate, Teaching Assistant Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Rok Fink, Bsc, MSc, Teaching Assistant Department of Sanitary Engineering, College for Health Studies, University of Ljubljana, Slovenia Sasa Erlih, BSc, MSc, Project Manager Institute for Comprehensive Development Solutions, Slovenia Ivan Erzen, MD, PhD, Associate Professor National Institute of Public Health and Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Ivan Erzen Chair of Public Health, Faculty of Medicine, University of Ljubljana, Zaloska cesta 4, 1000 Ljubljana, Slovenia e-mail: Ivan.Erzen@ivz-rs.si
Keywords	Air pollution, geographic information system, health effects.
Learning objectives	After completing this module students should: • understand why GIS is an important tool in the analysis of air pollution in environmental epidemiology, • be able to use AirGIS in practice with help of an instructor.
Abstract	Geographic Information System (GIS) is a software that encompasses storage, retrieval, analysis and display of spatial-geographical data. AirGIS is a human exposure modelling system for traffic air pollution. As a case study, AirGIS model was used in the case of air pollution in coast community Koper in Slovenia. Results of GIS are showing that wide circulation of asthma, chronically bronchitis and allergy is huge against children's account. This account is especially expressed in polluted areas. Results of analysis shows higher frequency of chronically lung diseases in more polluted than in less polluted local communities.
Teaching methods	Teaching methods include an introductory lecture, exercises, and interactive methods such as working on computer modelling. After the introductory lecture, students carefully read the recommended sources about traffic air pollution. Afterwards, they discuss air pollutants and practice on GIS model. At the end, students compare and discuss their results.
Specific recommendations for teachers	• ECTS: 0.2; • work in small group; • work proportion 50% / 50%; • facilities: a computer room; • reliable data for GIS study; • equipment: computers (1 computer on 2 students); • software: GIS; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple-choice questionnaire and case study.

THE GEOGRAPHIC INFORMATION SYSTEM (GIS): THE EXAMPLE OF TRAFFIC AIR POLLUTION

Andreja Kukec, Rok Fink, Sasa Erlih, Ivan Erzen

Theoretical background

Geographic Information System

About the Geographic Information System

Geographic Information System (GIS) is a software that encompasses storage, retrieval, analysis and display of spatial-geographical data.

GIS is a promising tool for exposure modelling due to the increase in coverage and quality of digital maps, developments in administrative databases managed by the authorities and developments of more user-friendly desktop GIS with increasing number of analytic features (1,2).

There exist different modules of this software, according to what the observed outcome is. For analyzing human exposure modelling systems for traffic air pollution, the AirGIS module is used.

About the AirGIS System

AirGIS is a human exposure modelling system for traffic air pollution. It was developed for application in Danish air pollution epidemiological studies, human exposure studies, as well as urban air quality assessment and management. It has been developed by the National Environmental Research Institute in Denmark. The model system allows for exposure estimates at a high spatial resolution (address level) and a high time resolution (hour). The system integrates air pollution dispersion models, digital maps, national and local administrative databases, concentrations of air pollutants at regional, urban background and street level, meteorological data, and GIS (3-5).

About using the Geographic Information System in analysis of traffic air pollution

Various exposure models have been developed that combine microenvironment concentrations with individual time-activity patterns and extrapolation to the entire population to give population exposure distributions (Figure 1). Traffic emissions are estimated based on emission factors together with average daily traffic (ADT), the percentage of heavy vehicles and the travel speed for each street section applying a default seasonal, weekly and diurnal variation in traffic loads to obtain hourly traffic inputs as well as default values for cold starts. The present project takes a microenvironment approach to population exposure modelling but adds a geographic dimension by taking advantage of GIS, digital maps and administrative databases. The objectives of the research project are: to develop an exposure model that combines modelled air pollution data using the Danish Operational Street Pollution Model (OSPM), population data using existing administrative databases, digital maps and GIS. A simple population dynamics model will be established to model the number of people present in a given area during a given time using simple profiles for time spent in the various areas at home, at work, and in transit. Additionally, ratios between indoor and outdoor concentrations will be taken into account (1,3,6).

Case study

Traffic air pollution and health concerns in Koper community/Slovenia

Introduction

Each and every one of our actions demands energy. Already early in history, man attempted to gain control over more energy, primarily through the use of animals and slaves. Not long after, we also learned through technical prowess to use nature's energy. It was only with Watts' invention of the steam engine in 1769 that it became possible for man to produce large amounts of energy on demand (7).

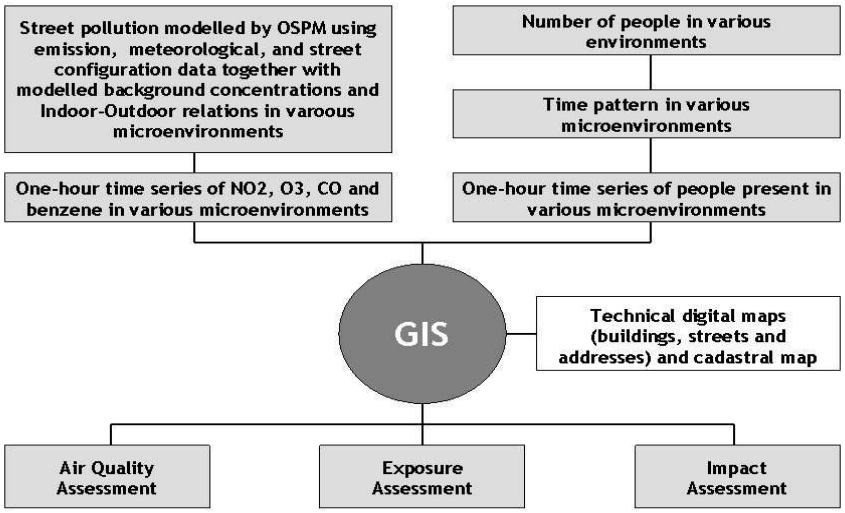
Transportation system has contributed significantly to the development of human civilization; on the other hand, it has an enormous impact on the air quality in several ways (8). In Western Europe, the transport of people and freight has dominated road traffic for many decades. Emissions from road traffic, from both combustion and friction processes, result in a complex mixture of air pollution, which is known to have adverse the population's exposure and the effects on health (8-10).

On average 1,4 persons are travelling by one car, therefore for transportation of 100 persons 70 cars are needed. From the point of environmental consideration, for 100 km 40 litters of gasoline are used, and with an average consumption of 8 litters per 100 km, this comprises 560 litters. Form every single litter of gasoline harmful compounds are emitted (11,12). Fuel combustion is the primary source of a large number of health-damaging air pollutants, including fine and respirable particulate matter ($PM_{2.5}$ and PM_{10}), carbon monoxide (CO), sulphur dioxide (SO_2), nitrogen oxides (NO_x), volatile organic compounds (VOCs), ozone (O_3), and atmospheric lead (13). Some of these pollutants are direct by-products of fuel combustion, but others (such as O_3) are formed in the air through chemical reactions with other agents in the atmosphere.

Nevertheless, world oil consumption is significantly rising. In year 2000, 76,6 million barrels per day were used, whereas seven years later (in 2007) there were used 85,7 million barrels per day. Trends showed that in 2009 almost 90 million barrels were used for transportation each day (14).

Of all the different types of pollution affecting human health, by far the most important one is air pollution (7). Human exposure is believed to cause severe health effects, especially in urban areas where pollution levels often are high. The classic example is the severe London smog (smoke and fog) episode in 1952 where the mortality rate in the city increased dramatically (8). Although transport emissions are rising faster in low-income and middle-income countries than in those with high income, there remain massive global inequalities in transport energy use both between and within countries (15).

Figure 1. Outline of the methodology for the exposure modelling (Adapted from Jensen, 1998) (1)



LEGEND: GIS- Geographic Information System; OSPM- Operational Street Pollution Model

Environmental effects of traffic air pollution

Combustion of liquid fossil fuels causes emissions due to its composition or type of combustion. Resulting pollutants are:

- **Sulphur oxides (SO_x):** Although SO_x is a symbol of all oxides of sulphur (e.g., SO_2 and SO_3), about 95% of all sulphur oxides are in the form of sulphur dioxide (SO_2). In the atmosphere, however, SO_2 is a precursor of highly destructive sulphates (SO_4^{2-}), formed by the chemical addition of oxygen (O_2). SO_3 is not a stable compound and may react with water (H_2O) to form sulphuric acid (H_2SO_4), a component of acid rain (16). Although road transport is a minor source of sulphur dioxide at the national level, in some urban areas it can be important. Raised concentrations of sulphur dioxide have been detected alongside busy roads (17).
- **Nitrogen oxides (NO_x):** Nitrogen oxides are powerful greenhouse gases with global warming potential 296 times higher than the carbon dioxide (18). They are formed during high-temperature combustion, largely from the nitrogen and oxygen present in the air, but also from the oxidation of nitrogen contained fuels. The main sources are internal combustion engines. Almost all NO_x are emitted as nitric oxide (NO) which is then rapidly oxidized to the more toxic nitrogen dioxide (NO_2) (17). During wintertime pollution episodes, NO_2 concentrations in urban areas may exceed internationally accepted air quality criteria set for the protection of human health. These cases were reported in some large industrial and urban population centres in the northwest Europe, particularly in the United Kingdom during stagnant wintertime weather conditions. During summertime pollution episodes, photochemical reactions driven by sunlight lead to the conversion of organic compounds and oxides of nitrogen into photochemical oxidants, in particular ozone. This phenomenon was first reported in Los Angeles in 1940's and has subsequently been observed in almost all urban population centres worldwide. These photochemical reactions also lead of the oxidation of sulphur dioxide (SO_2) into fine haze of sulphuric acid (H_2SO_4) (19).
- **Carbon oxide (CO):** Most anthropogenic carbon monoxide is generated in combustion processes. Internal combustion engines, both in on-road and in diverse off-road use, comprise a principal source. The majority of the carbon in automotive fuels is oxidized to carbon dioxide, while a small fraction is incompletely oxidized to carbon monoxide (17). Catalytic converters reduce CO emissions approximately by a factor of eight (7).

- **Volatile organic compounds (VOCs):** VOCs comprise a wide range of chemical compounds including hydrocarbons, oxygenates and halogen-containing species. However, a few VOCs are considered toxic to human health, such as benzene and 1,3 butadiene. The main sources of benzene are production, distribution and use of automotive fuels. Petrol vehicles emit more benzene than diesel vehicles, even when catalytic converters are used. Petrol and diesel contains little 1,3 butadiene, but is formed during combustion in the engine (17,20).
- **Particulate matter ($PM_{2,5\ 10}$):** PM is an air pollutant consisting of a mixture of solid and liquid particles suspended in the air. These suspended particles vary in size, composition and origin. The most commonly used size fractions are particles with an aerodynamic diameter between 2.5 μm and 10 μm. PM can either be directly emitted into the air (primary PM), or can be formed in the atmosphere from gaseous precursors (mainly sulphur dioxide, oxides of nitrogen, ammonia and non-methane volatile organic compounds). The most important chemical constituents of PM are sulphates, nitrates, ammonium, other inorganic ions such as Na^+, K^+, Ca^{2+}, Mg^{2+} and Cl^-, organic and elemental carbon, crustal material, particle-bound water and heavy metals (9). A major contribution of particulate matter in urban areas is believed to be attributable to emissions from diesel-powered vehicles and traffic. The fine and ultra-fine particles are generated mainly by combustion from diesel exhausts (13).
  Since 1950 the world population has more than doubled, and the global number of cars has increased by a factor of ten. In the same period, the fraction of people living in urban areas has increased by a factor of four (21). The environmental impact of transport depend on the transport volume (e.g. vehicle kilometres), technology, the way vehicles are used (speed, acceleration), the distribution of the use of vehicles over space and time, and the locations of exposed people, nature and buildings (22). But, rapid urbanization and increasing time spent in congested traffic means that exposure is increasing even where pollution levels are falling. The greatest burden is in the mega cities of developing countries (15). Due to Slovenian urbanization, green house gasses emission increased, especially CO_2. Results are showing that share of traffic transport with individual cars is rising; meanwhile public railway and bus transport are decreasing (23).
  Climate change is occurring in the context of increased anthropogenic stress across a range of natural systems including stratospheric ozone depletion, loss of biodiversity, spread of invasive species, exhaustion of wild fisheries, and the depletion of freshwater supplies. In addition to its own resource use, transport facilitates the exploitation of other resources. From a climate-change perspective the most important effect is opening up areas to deforestation. Through increasing the demands on our environment's carrying capacity, a capacity not known in advance - energy intensive transport - reduces its human carrying capacity (15,24).
- **Ozone (O_3)**: Ozone is a gas composed of three oxygen atoms. It is not usually emitted directly into the air, but at ground-level is created by a chemical reaction between oxides of nitrogen oxides and volatile organic compounds in the presence of sunlight. Ozone has the same chemical structure whether it occurs miles above the earth or at ground-level and can be "stratospheric" or "ground-level", depending on its location in the atmosphere. In the lower atmosphere of the Earth, ground-level ozone has a negative impact on environmental and human health. Motor vehicle exhaust and industrial emissions, gasoline vapours, and chemical solvents as well as natural sources emit NOx and VOC that help form ozone. Ground-level ozone is the primary constituent of smog. Sunlight and hot weather cause ground-level ozone to form in harmful concentrations in the air. As a result, it is known as a summertime air pollutant. Many urban areas tend to have high levels of ground - level ozone, but even rural areas are also subject to increased ozone levels because wind carries ozone and pollutants that form it hundreds of miles away from their original sources. Nevertheless, stratospheric ozone occurs naturally in the stratosphere approximately 10 to 30 miles above the earth's surface and forms a layer that protects life on earth from the sun's harmful rays (25).

Health effects of traffic air pollution

Air pollution has both acute and chronic effects on human health. Health effects range anywhere from minor irritation of eyes and the upper respiratory system to chronic respiratory disease, heart disease, lung cancer, and death.

Air pollution has been shown to cause acute respiratory infections in children and chronic bronchitis in adults. It has also been shown to worsen the condition of people with pre-existing heart or lung disease. Among asthmatics, air pollution has been shown to aggravate the frequency and severity of attacks. Both short-term and long-term exposures have also been linked with premature mortality and reduced life expectancy (13,26-28).

As we breathe, gases and particles of traffic exhaust are drawn into the lungs, where they contribute to a range of health problems. Pollutants can damage the lungs, as well as get into the bloodstream and travel to organs throughout the body (Figure 2) (29,30). When air pollution levels go up, there are more (29,30):

- emergency room visits,
- hospital admissions,
- asthma attacks,
- children absence from school
- deaths from lung and heart diseases.

Air pollution can affect both the respiratory and the cardiac systems. The health effects of air pollution can be seen as a pyramid (Figure 3), with the mildest but not common effects at the bottom of the pyramid, and the least common but more severe at the top of the pyramid. The pyramid demonstrates that as severity decreases the number of people affected increases (29,30).

Impacts of pollutants on health are as follows:

- **Sulfur oxides (SO_x):** Sulphates and sulphur-containing acids are more toxic than sulphur dioxide gas. They interfere with normal functioning of the mucous membrane in respiratory passages, increasing susceptibility to infection. At high concentration levels, both H_2S and organic sulphur compounds are toxic (16). Sulphur dioxide is a water soluble irritant gas (28). Sometimes causing discomfort and coughing in healthy people, and often causing severe respiratory symptoms in asthmatics. When asthmatics were exposed under controlled conditions to levels of sulphur dioxide similar to those found near pollution sources such as ports, they developed an average decrease of 25–30% in their lung function (26). However, on occasional down-drafting of smokestack plumes or meteorological inversions near point-sources result in low parts per million (ppm) levels of SO_2 that may pose a hazard to some individuals. A two-minute exposure to 0.4 to 1.0 ppm can elicit bronchoconstriction in exercising asthmatics within 5 to 10 min (16). Once deposited along the airway, SO_2 dissolves into surface lining fluid as sulphite or bisulphite and is readily distributed throughout the body. It is thought that the sulphite interacts with sensory receptors in the airways to initiate local and centrally mediated bronchoconstriction (28).
- **Nitrogen oxides (NO_x):** Numerous studies have found that NO_x can cause toxic effects on the airways, leading to inflammation and to asthmatic reactions. In fact, people with allergies or asthma have far stronger reactions to common allergens such as pollens when they are also exposed to NO_x (26). In health effects, NO_2 can irritate the lungs and lower resistance to respiratory infections (31). Persons suffering from respiratory diseases, such as asthma, are very sensitive to NO_2 at high concentrations, while several risk assessment studies have shown that both short and long term exposure to NO_2 can induce effects to the human health and that, given the role of NO_2 as a precursor of other pollutants and as a marker of traffic related pollution, there should be benefits for the public health from keeping low NO_2 levels in the atmospheric air. Recent studies have shown that chronic exposure to the levels of air pollutants, such as NO_2, currently observed may have even higher impacts on mortality than acute exposure (32).
  Several studies indicate that the combination of SO_x and NO_x in the air is particularly noxious, because these compounds appear to act together to increase allergic responses to common allergens such as pollen and dust mites (26).
- **Carbon monoxide (CO):** Carbon monoxide is classed toxicologically as a chemical asphyxiant because its toxic action stems from its formation of carboxyhemoglobin, preventing oxygenation of the blood for systemic transport (31). Carbon monoxide reduces the ability of blood to deliver oxygen to the cardiovascular and nervous systems. Long-term exposure can cause brain damage due to the lack of oxygen reaching the brain. The symptoms of carbon monoxide exposure are varied and include dizziness, nausea, fatigue, and decreased muscular control (33).
- **Volatile organic compounds (VOCs):** Benzene and butadiene are known to cause cancer in humans. Formaldehyde is very irritating to the airways, and is a probable carcinogen. Toluene at occupational exposure levels has been associated with birth defects and miscarriages. Other VOCs emitted by vehicles have also been linked to cancer, reproductive harm, asthma, or neurological disorders (26).
- **Particulate matter ($PM_{2,5\ 10}$):** Fine particles are believed to be the most critical pollutant and the main contributor to excess mortality especially among people suffering from respiratory and cardiovascular diseases (1). Elevated PM_{10} concentrations have been linked with increases in mortality and hospital admissions with respiratory symptoms, especially for sensitive population subgroups (9,34). Air pollution is a major environmental health problem causing approximately three million deaths per year in the world, as a result of exposure to particulate matter (35). Correlation between increased exacerbation of respiratory diseases, cardiopulmonary morbidity, mortality and the levels of urban airborne particulate matter is now well-established (36).
- **Ozone (O_3)**: Exposure to ozone as a secondary component formed in the atmosphere at some distance from traffic is widespread and not very different for people living in rural and urban areas, although it depends largely on the time spent outdoors (26). Thousands of scientific studies have been published on the health effects of ozone. Ozone can make people more susceptible to respiratory infections and can aggravate pre-existing respiratory diseases, such as asthma. Ozone can also cause irreversible changes in

the lung structure, which eventually lead to chronic respiratory illnesses, such as emphysema and chronic bronchitis (26,31).

Figure 2. Air pollution effects in the body (Adapted from Mishra, 2003) (30)

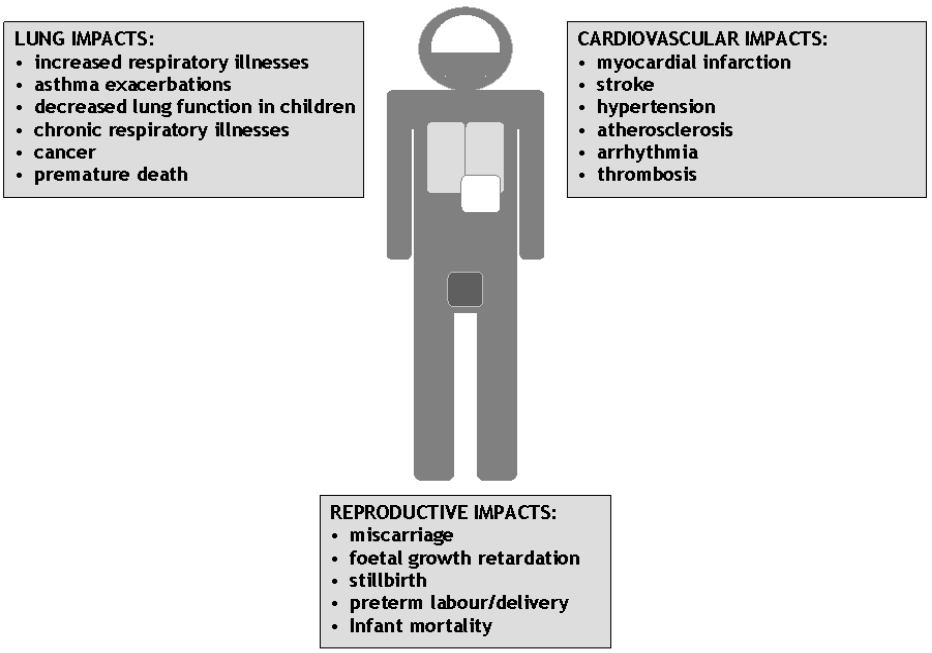


Figure 3. Pyramid of health effects of air pollution (Adapted from Mishra, 2003) (30)

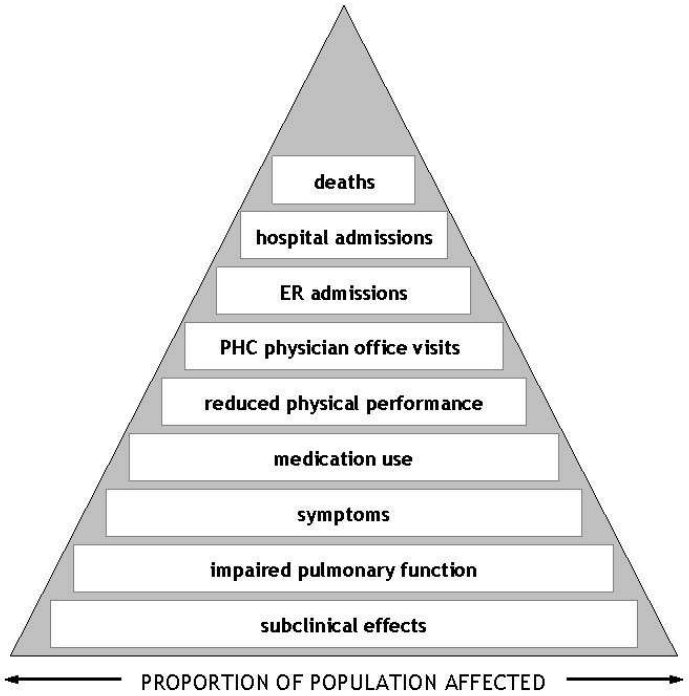
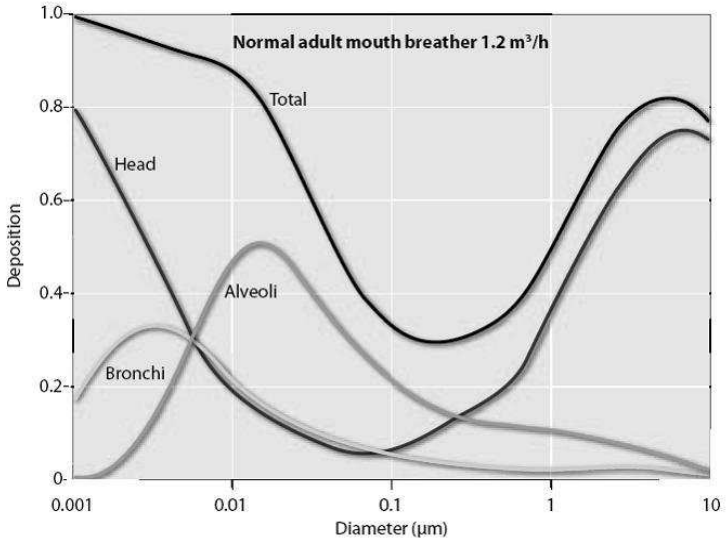


Figure 4 shows where particles are deposited in the respiratory tract, depending on their size. Smaller particles (in particular, $PM_{2.5}$) penetrate more deeply into the lungs and may reach the alveolar region. Ultra fine particles contribute only slightly to PM_{10} mass, but may be important from a health point of view because of the large numbers and the high surface area. They are produced in large numbers by combustion, especially the internal combustion engines (29,30).

Figure 4. Deposition probability of inhaled particles in the respiratory tract according to particle size (Adapted from Mishra, 2003) (30)



Exposure and Health Effects

Exposure estimates to atmospheric pollutants can address individuals (personal exposure) or large population groups (population exposure), and can be based on direct (exposure monitoring) or indirect methods like exposure modelling (35). Exposure to ambient air pollution has been linked to a number of different health outcomes, starting from modest transient changes in the respiratory tract and impaired pulmonary function, continuing to restricted activity/reduced performance, emergency room visits and hospital admissions and to mortality. There is also increasing evidence for adverse effects of air pollution not only on the respiratory system, but also on the cardiovascular system. This evidence stems from studies on both acute and chronic exposure (Table 1) (9).

Table 1. Important health effects associated with exposure to different air pollutants (9)

Pollutant	Effects related to short-term exposure	Effects related to long-term exposure
1. Particulate matter	• lung inflammatory reactions • respiratory symptoms • adverse effects on the cardiovascular system • increase in hospital admissions • increase in mortality	• increase in lower respiratory symptoms • reduction in lung function in children • increase in chronic obstructive pulmonary disease • reduction in lung function in adults • reduction in life expectancy, owing mainly to cardiopulmonary mortality and probably to lung cancer
2. Ozone	• adverse effects on pulmonary function • lung inflammatory reactions • adverse effects on respiratory symptoms • increase in medication usage • increase in hospital admissions • increase in mortality	• reduction in lung function development
3. Nitrogen dioxide	• effects on pulmonary function, particularly in asthmatics • increase in airway allergic inflammatory reactions • increase in hospital admissions • increase in mortality	• reduction in lung function • increased probability of respiratory symptoms

Some groups may receive much higher exposures than others, such as people who live and work near busy roads and those who travel or commute in heavy traffic. Also, the intake of pollutants by road users varies between drivers, bicyclists and pedestrians. Further, exposure to transport-related air pollution is difficult to separate from exposure to total air pollution (9,27,28).

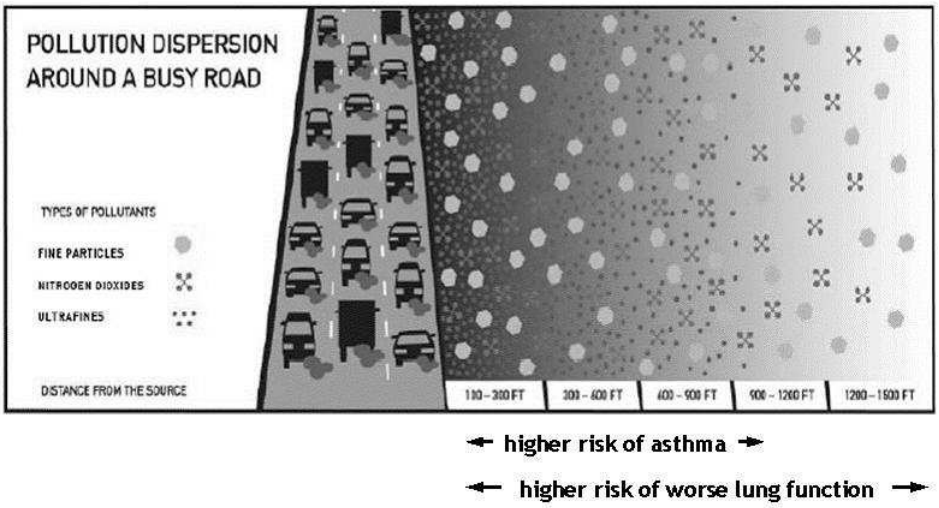
People on road transport, in residencies and schools near the main roads are exposed to high levels of vehicle pollutant emissions (13).

Population exposure to traffic pollution can be classified into three general categories:

- contribution of traffic emissions to rural and urban background air pollution;
- locally elevated levels of traffic pollution near busy roadways and city zones; and
- still further elevated exposure concentrations in commuting/transport (29,31).

In Figure 5 are presented three types of pollutants and their impact range.

Figure 5. Traffic spreads pollution up to 1500 feet (about 450 meters) from the roadway (Adapted from Yeh, 2005) (31). NOTE: 1 foot equals to 0.3048 m. (LEGEND: higher risk of asthma up to 500 feet (160 m); higher risk of worse lung function)



Geographical dispose of chronically lung diseases at children in Koper community

Introduction

Chronically, lung diseases are becoming an important public health concern mainly in developed countries. Asthma, chronically bronchitis and allergy are in many cases present in childhood. The most frequent lung illness among children in the developed world and in Slovenia is asthma (37).

Inhabitants in the west and north parts of Koper community were complaining about lung diseases. How so ever, analyses showed that mortality in Koper community in case of lung illnesses is higher than in others Slovenian communities. In the year 2003 study, an improved focus on health in different geographical parts of Koper community was made. The aim of this study was to found out if health risk is higher in parts were harmful factors exist compared to other areas where these environmental factors are less present (37).

Geographical review

Data

Different sources of data are used in GIS system:

1. For research purposes geographical information system was used, to show lung illness distribution. In the system mentioned above data were collected, saved, controlled, processed and analyzed.
2. Two kinds of data can be joined in GIS; descriptive and geographical attributes. From the poll that contained address for each questionnaire, data about location or geocode (X and Y coordinate; Gauss-Krueger coordinate system D48/GK) were made. Relation base was made from existing data; number of questionnaire, geocode of location (X, Y), distribution regarding community area, distribution regarding postal codes, and distribution regarding health status.
3. Data were used to calculate prevalence of chronic diseases: asthma, chronic bronchitis and allergy. Nevertheless, another base of illnesses that is in GIS connected with geographically survey of different data groups was made (37).
4. Data base (vector data) was moved in GIS, so that questionnaires could be located. Cartographically pads of national topographically maps in measure 1:25.000 and 1:50.000 (DTK

25 and DTK 50) were used (raster presentation). Maps are in the ownership of surveying and mapping authority of the Republic of Slovenia. In GIS was also incorporated artificial personal data that due to activity can influence on environment. Data were collected in a special study, which was conducted by the company OIKOS d.o.o. in the year 2006.

5. In GIS, system layers were entered; borders of Koper community, borders of local communities. For process and review of data program, the software ESRI Arc View 3.2 1996 was used (37).

Air pollution

Industry, traffic and combustion devices are the most significant sources of air pollution in Koper community. Whole year sources are industry and traffic. Meanwhile, weekly and seasonal dynamic is significant for traffic pollution. The higher emissions are present in summer and at the end of the week, when traffic is the most dense; nevertheless pollution from combustion devices is important only in winter.

Regulation and air pollution control

The most direct means of improving air quality is of course through regulation of emissions. The EU legislation on vehicle emission and fuel quality standards has evolved greatly since the first directive in 1970. The early legislation had the dual purpose of reducing pollution and avoiding barriers to trade due to different standards in different member states.

The "Auto Oil I Programme" undertaken by the European Commission in conjunction with the industry, has set up the targets for a series of traffic-related pollutants and assessed different technologies and fuel quality standards (38,39). The base for air quality regulation in Slovenia is the Environment Protecting Act (RS) No. 41/2004, 17/2006, 20/2006, 28/2006, 39/2006, 49/2006, 66/2006, 112/2006, 33/2006 (40). However, in the year 2005, Slovenian government accepted the Decree on Nation Emission Ceilings for Atmospheric Pollutants (RS) No. 24/2005, 92/2005, where the national maximum emissions of SO_2, NO_x, VOC and NH_3 were established (41). Strategies and arrangements for the implementation of this decree are described in the operational programme on national emission ceilings for atmospheric pollutants in ambient air (RS) No. 24/2005 (42). Some specific claims for each individual pollutant is defined in the Decree on Benzene and Carbon Monoxide in Ambient Air (RS) No. 52/2002 and the Decree on Sulphur Dioxide, Nitrogen Oxides, Particulate Matter and Lead in Ambient Air (RS) No. 52/2002, 18/2003, 121/2006 (43,44).

Table 2. Maximum concentrations of some pollutants (WHO-Air quality guidelines) (45)

Pollutant	Time interval	Maximum concentration
PM 2,5	Annual	10 µg/m^3
	24-hour	25 µg/m^3
PM 10	Annual	20 mg/m^3
	24-hour	50 mg/m^3
O3	8-hour	100 µg/m^3
NO2	Annual	40 µg/m^3
	1-hour	200 µg/m^3
SO2	24-hour	20 µg/m^3
	10-minute	500 µg/m^3

Practical work with GIS

The process of working with the GIS model could be split into several steps. The addresses were converted into spatial location data, which is needed to present, analyze, and attribute data in spatial settings. We used point data attained from geocoding and regional data to give an overview of the situation (diminish variability in population counts for regions).

Data insertion and processing

Geocoding

Geocoding is used as a method for assigning a spatial location to an address record. Although many possibilities exist, we used public available databases of Environmental atlas of Slovenia. Point shape files were created:

1. Locations information (local community, street and number) were collected with poll in a pilot research "Environmental impact on incidence of some illness and mortality in Koper community" in the time period 2002-2003.
2. Since information about permanent residence is indirect data in each poll, data about location (geocoded address) are in two different types:
    - Environmental atlas in the website of Environmental Agency of the Republic of Slovenia (http://gis.arso.gov.si/atlasokolja/), where a single person can be determined based on the

permanent address. Program house numbers can be found with coordinates X and Y and resolution of 1: 1500. Data are then copied in the database (C/P).

- Also, for a better matching and indirect precision, it was performed a control of geolocation test with data base ENHIS (evidence of house numbers, dBase, GURS).

Geocodes can be checked up in case of every single address. In the birth database, matching local community can be found.

Relation database

Result of geocoding is the relation database (contains the following data; poll cipher, geocoded location (X, Y), local community, post number and health condition) in Microsoft Excel and then transformed into dBase for program ESRI ArcView 3.2 and new version of ESRI ArcMap 9.2.

Data insertion in GIS

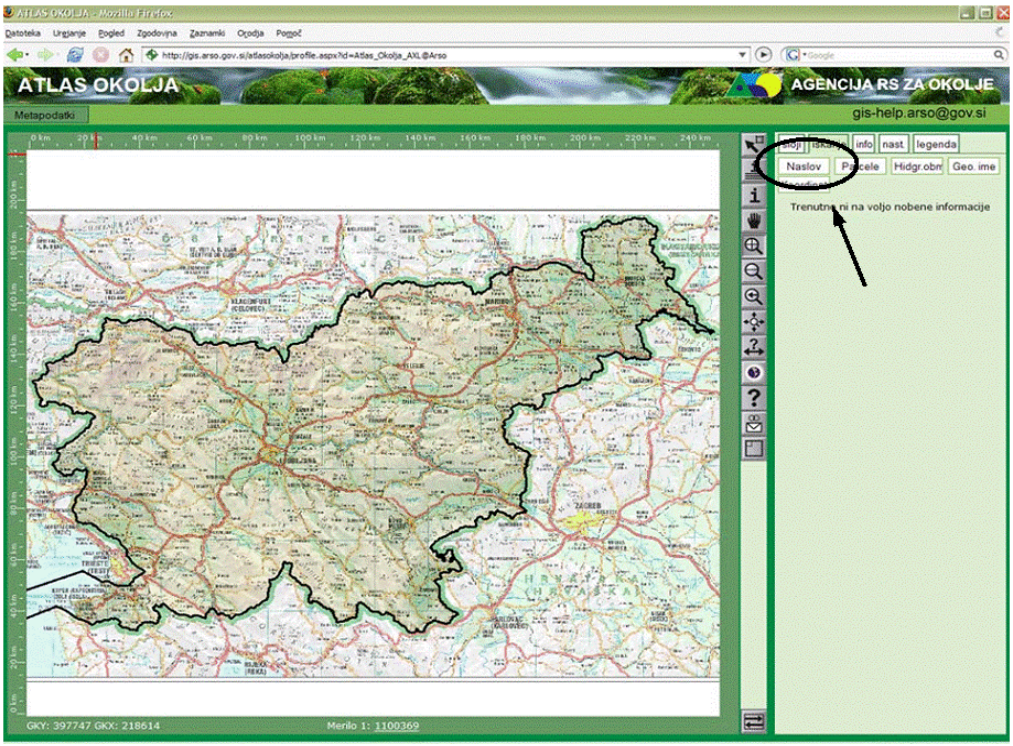
In GIS, the sphere geolocated raster map (.tiff) is an adequate measure. Data that has to be in dBase are transformed in individual layers. Further, data manipulation like community borders and addition of others databases is dependent from program packages.

Algorithm

1. Location control:
   In the program, choose "Search" and then "Parcel address". Insert community and settlement. Figure 6 shows where in the communication window to insert address in Environmental Atlas (Figure 6). Geocoding address can be more easily carried out through the base ENHIS. The problem may be represented by the availability of databases, or its price. Given that, the basis for search in the Environmental Atlas is better.
   In continuation, find street and number of street. The program will show the exact location. Figure 7 shows where to choose a command to run procedure for information about coordinates (Figure 7).
   Mark both coordinates and transform them in Microsoft World or Excel.
2. Address control with database of birth (91–96) and community, post number. Every single address, cipher in controlled with database of birth 91–96.
3. Final database (X, Y coordinates, community and post number). It is simple to transform data.
4. Insertion of raster map (1:50.000) in GIS (ArcMap).
5. Insertion of database with definition of data fields (point feature or area feature).

Figure 6. Where to insert address in Environmental Atlas

a)



b)

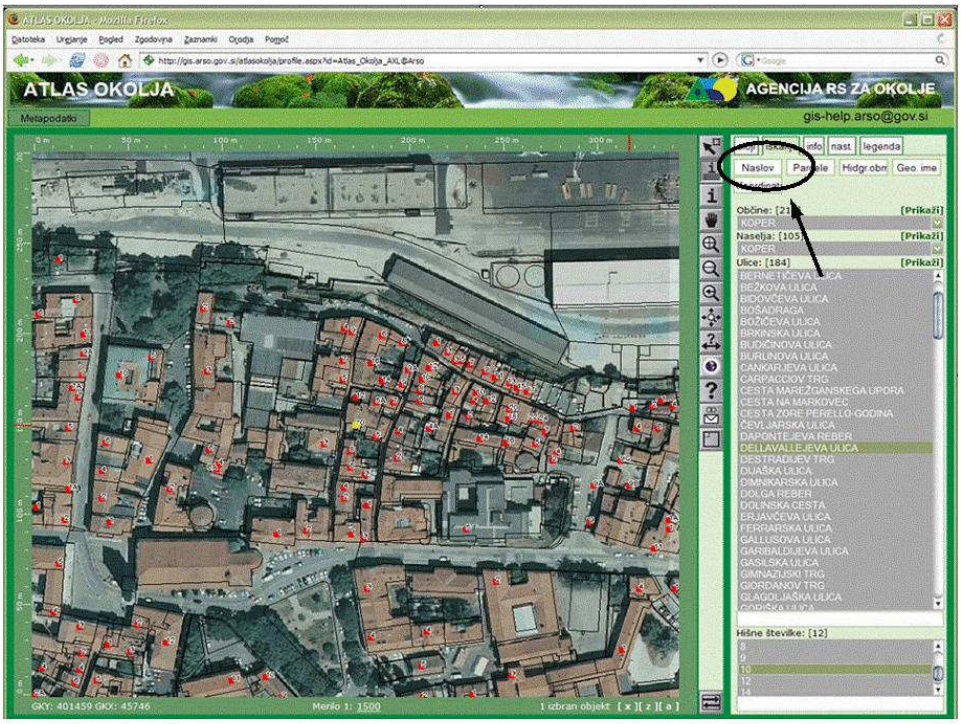


Figure 7. Where to choose "i" for information about coordinates

Figure 8. Function File/Add data recognize TIFF data for location of map in space. In the left field, the title of map can be seen

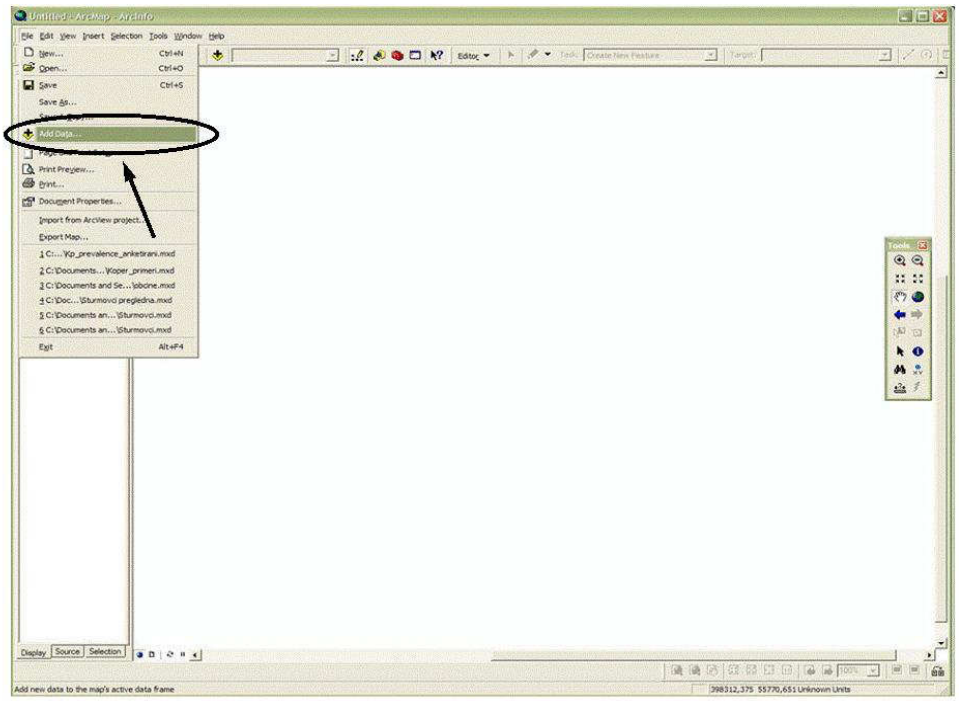


Figure 9. Coordinates X, Y are transformed in ArcMap 9.2 with simple function were only names of colons are necessary (Function Tools/Add X, Y data). Choose base /shape

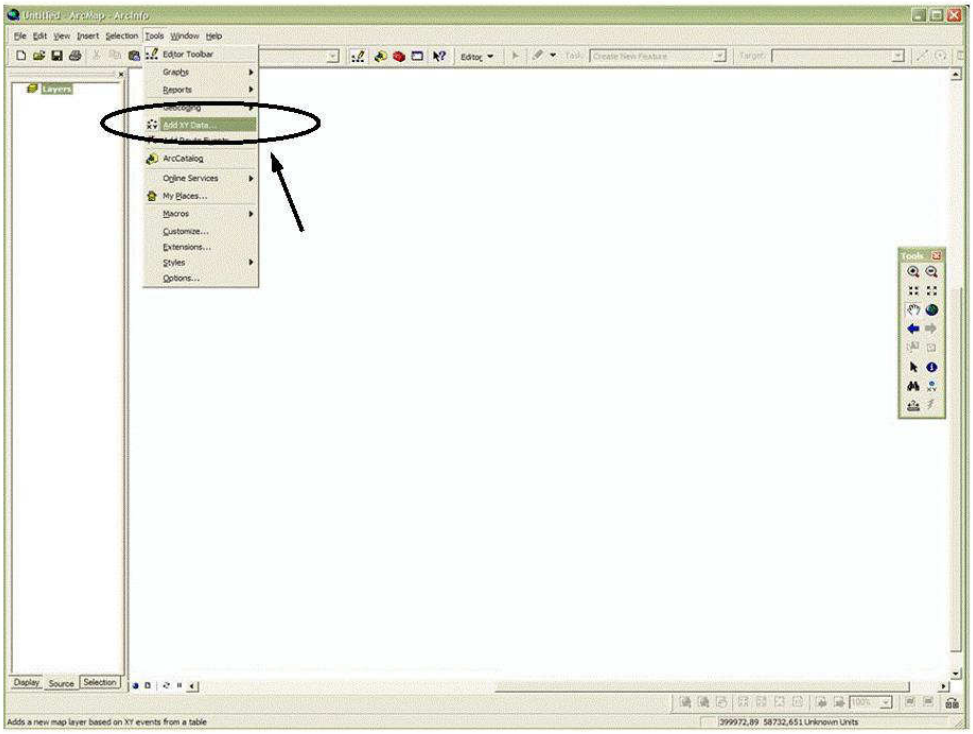
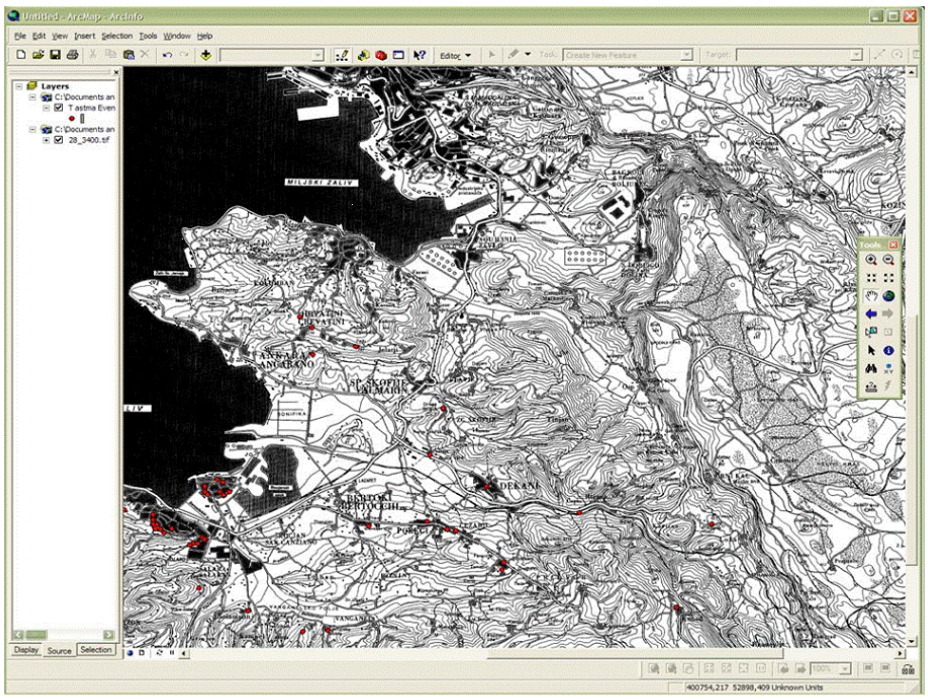


Figure 10. The basic view of coordinates can be seen. With right mouse button individual properties, correlations, changing properties can be done



Algorithm of assigning areal or regional data

1. An Excel table for summing addresses in local community is prepared. This table is compared with the table of birth (1991-1996).
2. Prevalence in the community, global prevalence and comparisons are made (in Excel).
3. In sphere GIS "GIS shape" communities (two dimensional form of community borders) are transferred.
4. With function Add data:
    - in sphere GIS table need to be complement ("Attribute table" and function "Edit". Changeability, prevalence and constant risk hypothesis, etc.).
    - data manipulation (clusters of changeability, prevalence and different levels of review).
    - covering of different layers (Hypothesis about correlation of sea level can be check up. Sea level in insert in GIS and adjacency of single case with isohyps. Histogram of cases in isohyps is result).

On the base of information in "view" is it difficult to make final correlation about illness cases. This is especially important in small areas with relatively small numbers of cases. Only a quick look on unworked maps will give you an answer why to use "smoothing" like the one of the view possibility.

Appearance overview of chronic lung diseases

An overview of all chronic lung diseases shows accumulation of cases in several local communities (Figure 6). The risk for some chronically lung illnesses in childhood if compared with control community Smarje pri Jelsah is higher in the local community Pobegi Cezarji (4,23 times) and Hrvatini (3,67 times). In this area, a high prevalence of lung illnesses is present also. Due to data in local community Pobegi Cezarji 31,25% children in the fourth class of primary school have chronic lung disease. In case of Hrvatini, 27% of the children suffer for the same disease (37).

Geographic distribution of children' addresses

An overview of the geographic distribution of children's addresses that has chronic lung disease showed that some areas are more risky for lung illnesses than others. Significant accumulation can be seen in local areas of Pobegi Cezarji, Hrvatini, Olmo Prisoje, Zagrad. The relatively high prevalence in Merezige, Vaganel, Dekani and Crni Kal must be interpreted carefully due to the low levels of cases, as distribution can be occasional (Figure 11-15).

More specifically, the overview shows that children with chronic lung disease live in higher areas that are located in the nearby slope. Usually, a permanent address is located 700 and 900 meters above the sea level (37).

Figure 11. Geographic distribution of permanent addresses of children that were in 2002/2003 in the fourth class of the primary school and had asthma. Distribution by local community (37)

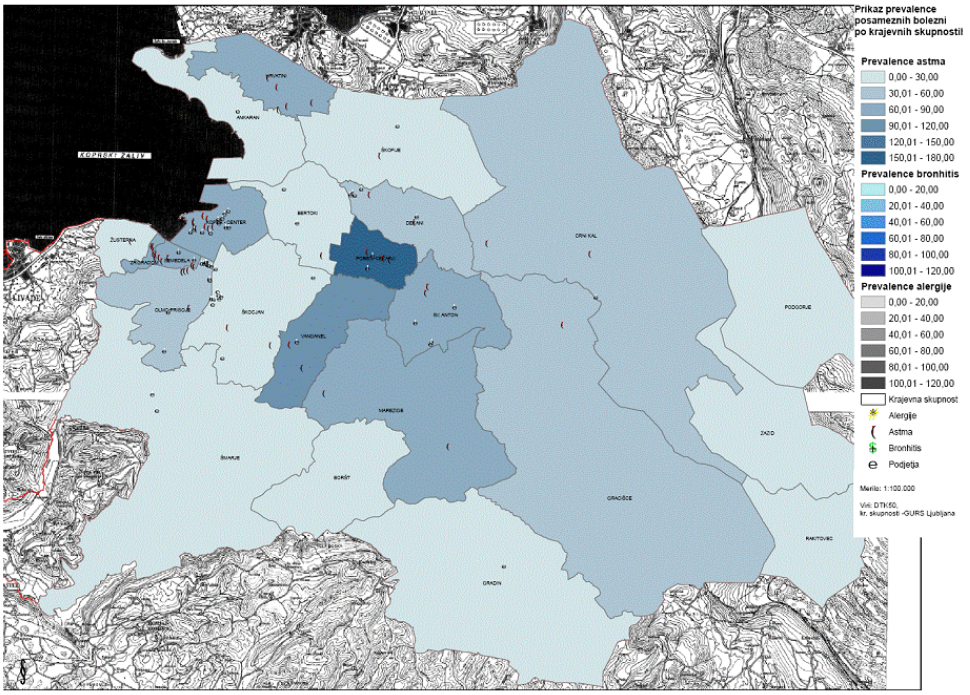


Figure 12. Geographic distribution of permanent addresses of children that were in 2002/2003 in the fourth class of the primary school and had chronic bronchitis. Distribution by local community (37)

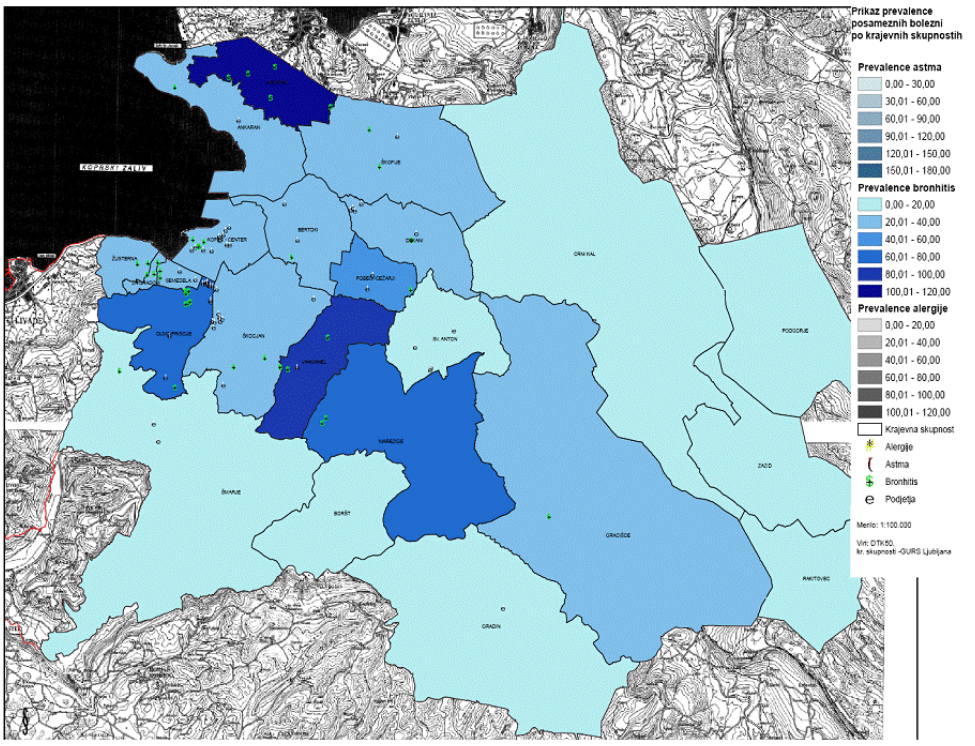


Figure 13. Geographic distribution of permanent addresses of children that were in 2002/2003 in the fourth class of the primary school and had asthma or chronic bronchitis. Distribution by local community (37)

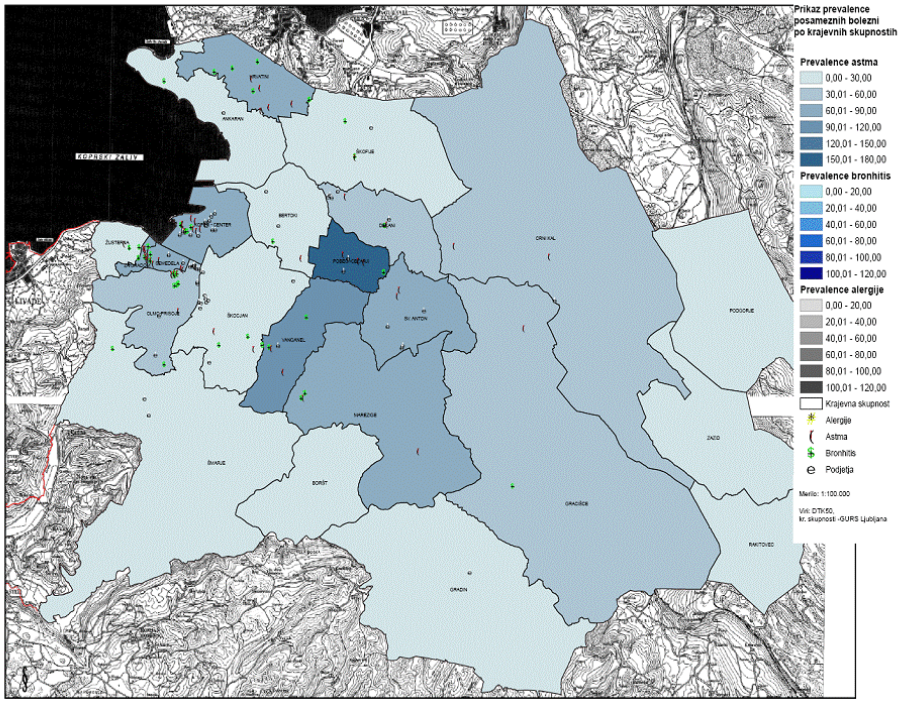


Figure 14. Geographica distribution of permanent addresses of children that were in 2002/2003 in the fourth class of the primary school and had lung allergy. Distribution by local community (37)

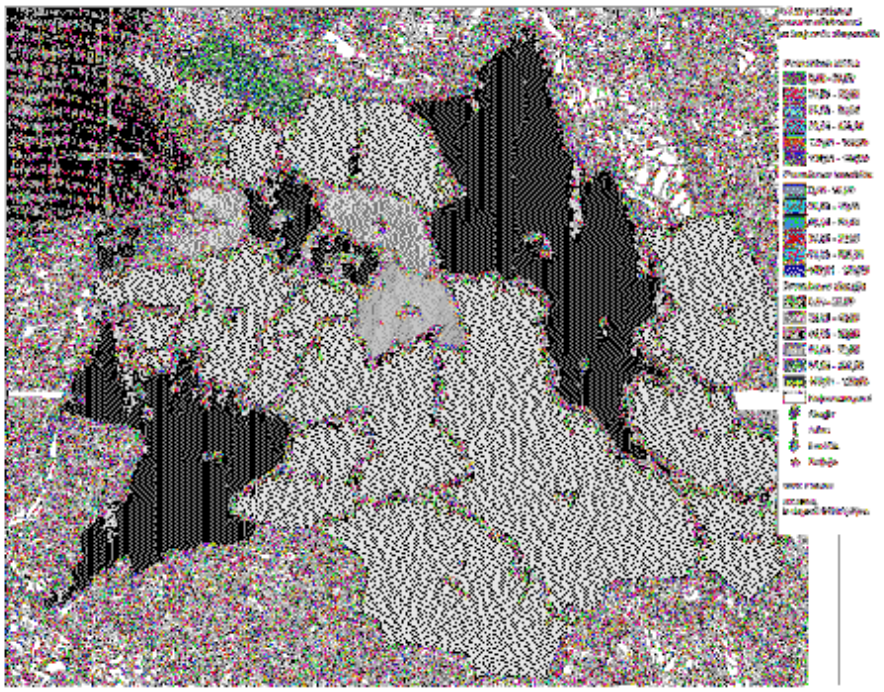
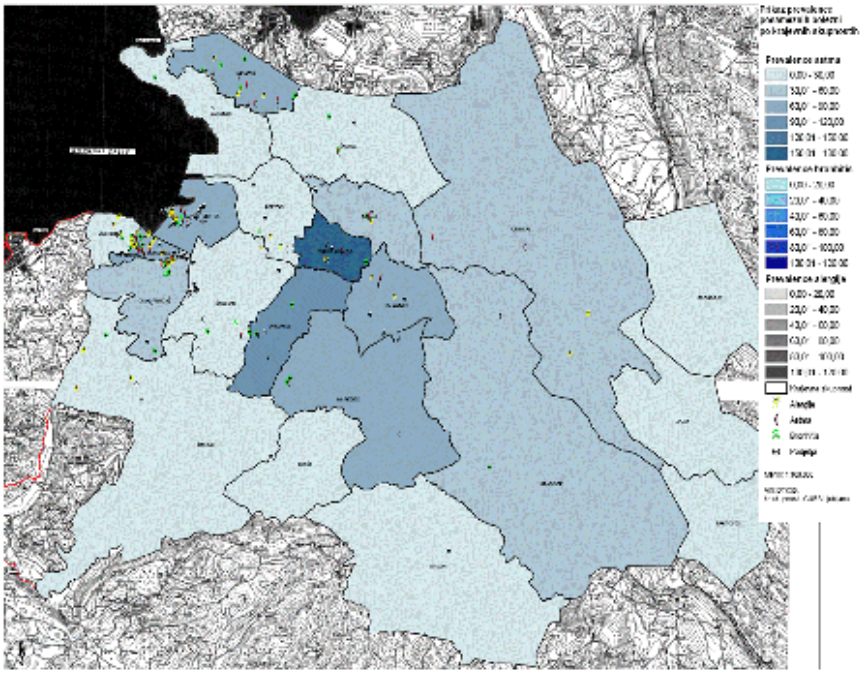


Figure 15. Geographic distribution of permanent addresses of children that were in 2002/2003 in the fourth class of the primary school and had chronic lung disease. Distribution by local community (37)



Conclusion

Results of this study show that wide circulation of asthma, chronic bronchitis and allergy are huge against children's account. This account is especially expressed in polluted areas. Results of analysis shows higher frequency of chronic lung diseases in more polluted than in less polluted local communities.

At this stage, it is impossible to determine which pollution sources are the most important ones for chronic lung illnesses. With no doubt, the problem is very complex.

In the area of Koper community, further studies about pollution sources and their impact on health should be undertaken. Nevertheless, efforts for pollution decreasing irrespective if this is correlated with health problems or not should be immediately undertaken. In regard to this conclusion, the traffic air pollution is the first task to be performed.

Exercises

Task 1

- Carefully read the theoretical background of this module and discuss with other students about it.
- Traffic air pollution and pulmonary diseases: Which pollutants have the major impact on human health?
- Can pollutants be reduced? What is the problem of reducing them?
- GIS model as a tool in epidemiology. What are the advantages and disadvantages?

Task 2

Choose one pulmonary disease and insert disease data in GIS programme.

Task 3

Compare your results with some other models used in epidemiology.

References

1. Jensen SS. Mapping human exposure to traffic air pollution using GIS. Journal of Hazardous Materials 1998;61:392-385.
2. Gulliver J, Briggs JD. Time-space modeling of journey-time exposure to traffic-related air pollution using GIS. Environmental Research 2005;97:25-10.
3. Cromley EK, McLafferty SL. GIS and Public Health. New York, 2002.
4. National Environmental Research Institute. AIRGIS. Available from: URL: http://www.dmu.dk/International/Air/Models/AIRGIS/ (Accessed: 15 September 2013).

5. Tang UW, Wang ZS. Influences of urban forms on traffic-induced noise and air pollution: Results from a modelling system. Environmental Modelling & Software 2007;22:1764-1750.
6. Jarup L. The role of geographical studies in risk assessment. In Spatial Epidemiology. Methods and Applications, 2006:432-412.
7. Lomborg B. The skeptical environmentalist. Measuring the real state of the world. Cambridge university press 2001:515.
8. Mahboob A, Makshoof A. Air pollution due to traffic, air quality monitoring along three sections of National Highway N-5, Pakistan. Environ Monit Assess 2008;136:226-219.
9. World health organization (WHO). Health risk of particulate matter from long-range transboundary, 2006. Available at URL: http://www.euro.who.int/document/E88189.pdf.
10. World health organization (WHO). Health aspects of air pollution results from the WHO project "Systematic review of health aspects of air pollution in Europe". June, 2004. Available from: URL: http://www.euro.who.int/document/E83080.pdf.
11. Ostro B. World health organization (WHO). Outdoor air pollution. Genava: 2004. Available at URL: http://www.who.int/quantifying_ehimpacts/publications/ebd5.pdf (Accessed: 15 September 2013.).
12. Lah A. Remodeling of transport systems (in Slovene). In: Traffic and environment. Collection adjustment and sustainable. Environment protection world Republic of Slovenia, 2002.
13. Kongtip P, Thongsuk W, Yoosook W, Chantanakul S. Health effects of metropolitan traffic-related air pollutants on street vendors. Atmospheric Environment 2006;40:7145-7138.
14. Energy Information Agency (EIA). Official Energy Statistic from U.S. Government, 2008. Available at URL: http://www.eia.doe.gov/oil_gas/petroleum/info_glance/petroleum.html (Accessed: 15 September 2013).
15. Woodcock J, Banister D, Edwards P, Prentice A M, Roberts I. Energy and transport in Energy and health 2007;3:1078-88.
16. Wang LK, Williford C, Chen WY. Desulphurization and Emissions Control. In Advanced air and noise pollution control 2005:92-35.
17. Holman C. Sources of air pollution. In: Holgate S, Samet JM, Koren HS, Maynard RL, et al. Air pollution and health. London: Academic Press, 1999:115–48.
18. Fogg PGT, Sangster JM. Chemicals in the Atmosphere- Solubility, Sources and Reactivity. Wiley, 2003: 7.
19. Dervent RG. Atmospheric Chemistry. In Air pollution and health. Holgate ST, Samet JM, Koren SH, Maynard RL. London: Academic Press, 1999:63-51.
20. Colls J. Air Pollution. 2th. Ed. London: Spon Press, 2002.
21. Fenger J. Urban air quality. Atmospheric Environment 1999;33(29):4900-4877.
22. Garling T, Steg L. Environmental effects of urban traffic. Threats from Car Traffic. The Quality of urban life. Problems, Causes, Solutions. 1 edition. Elsevier Science 2007:33-11.
23. Bernard-Vukadin B et al. Environment on hand. Sustainable consumption and production (in Slovene). Edit by Rejc-Brancelj I, Zupan. Environmental Agency of the Republic of Slovenia (EARS) 2007:89. Available at URL: http://nfp-si.eionet.europa.eu/Podatki_in_informacije/F1203936512/HTML_Page1203951969 (Accessed: 14 March 2013).
24. Schultz MG, Feichter J, Leonardi J. Climatic impact of Surface Transport. In Transport and the Environment. Issues in Environmental Science and Technology. Royal Society of Chemistry 2004:126-111.
25. EPA (Environmental protection agency). Ground-level Ozone, 2008. Available at URL: http://www.epa.gov/air/ozonepollution/ (Accessed: 14 March 2013).
26. Bailey D, Solomon G. Pollution prevention at ports: clearing the air. Environmental Impact Assessment Review 2004;24:774–749.
27. Schnelle B K, Brown AC. Air pollution control technology handbook. CRC Press, 2001.
28. Wright DA, Welbourn P. Environmental toxicology. Cambridge: Cambridge University press, 2002.
29. Health & Environment Impact Project, "Moving Forward" Conference 2007, University of Southern California, 2007 Available from: URL: http://www.scribd.com/doc/907226/2-Air-Pollution-Health-Effects-2007 (Accessed: 15 February 2013).

30. Mishra V. Health Effects of Air Pollution, 2003. Available from: URL: http://www.mnforsustain.org/climate_health_effects_of_air_pollution_mishra_pern.htm (Accessed: 15 February 2013).
31. Yeh TJ, Chen WY. Control of NOx During Stationary Combustion. In Advanced air and noise pollution control 2005:124-113.
32. Chaloulakou A, Mavroidis I, Gavriil I. Compliance with the annual NO2 air quality standard in Athens. Required NOx levels and expected health implications. Atmospheric Environment 2008; 42: 265-454.
33. New Hampshire Department of Environmental Services (NHDES). Air Resources, 2006. Available from: URL: http://www.des.state.nh.us/ard/polut.htm (Accessed: 15 February 2013).
34. Grivas G, Chaloulakou A, Kassomenos P. An overview of the PM10 pollution problem, in the Metropolitan Area of Athens, Greece. Assessment of controlling factors and potential impact of long range transport. Science of the total environment 2008;389:177-165.
35. Borrego C, Tchepel O, Costa A M, Martins H, Ferreira J. Urban Population Exposure to Particulate Air Pollution Induced by Road Transport, 2007.
36. Garçon G, Dagher Z et. al. Dunkerque City air pollution particulate matter-induced cytotoxicity, oxidative stress and occurrence of human epithelial lung cells (L132) in culture. Journal of Toxicology in vitro 2006;20:528-519.
37. Erzen I, Zaletel-Kragelj L, Erlih S. Examining the characteristics of the incidence of chronic respiratory diseases in the area of the Municipality of Koper in the light of the potential Aetiological factors
38. Fenger J. Urban-Scale Air pollution. In. Handbook of Atmospheric Science. Principle and Applications. Blackwell Publishing, 2003:438-401.
39. Environment Protecting Act (RS) No. 41/2004, 17/2006, 20/2006, 28/2006, 39/2006, 49/2006, 66/2006, 112/2006, 33/2006
40. Blanken PD, Dillon J, Wismann G. The impact of an air quality advisory program on voluntary mobile source air pollution reduction. Atmospheric Environment 2001;35:2421-2417
41. Decree on sulphur dioxide, nitrogen oxides, particulate matter and lead in ambient air (RS) No. 52/2002, 18/2003, 121/2006.
42. Decree on nation emission ceilings for atmospheric pollutants (RS) No. 24/2005, 92/2005.
43. Decree on benzene and carbon monoxide in ambient air (RS) No. 52/2002. Operational programme on national emission ceilings for atmospheric pollutants in ambient air (RS) No. 24/2005.
44. Decree on sulphur dioxide, nitrogen oxides, particulate matter and lead in ambient air (RS) No. 52/2002, 18/2003, 121/2006.
45. World health organization (WHO). WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide, 2005. Available from: URL: http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf (Accessed: 15 February 2013).

Recommended readings

1. Jarup L. The role of geographical studies in risk assessment. In Spatial Epidemiology. Methods and Applications, 2006:432-412
2. Jensen SS. Mapping human exposure to traffic air pollution using GIS. Journal of Hazardous Materials 1998;61:392-385.
3. National Environmental Research Institute. AIRGIS. Available from: URL:
4. http://www.dmu.dk/International/Air/Models/AIRGIS/ (Accessed: 15 February 2013).

HEALTH INVESTIGATION: ANALYSIS – PLANNING – EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Basic occupational health indicators on sick leave**
Module: 2.21	ECTS (suggested): 0.15
Author(s), degrees, institution(s)	Marjan Bilban, MD, PhD, Professor, Head of the Chair Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Marjan Bilban Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: marjan.bilban@zvd.si
Keywords	Health indicators, occupational health.
Learning objectives	After completing this module students should: • know the definition and characteristics of occupational health indicators on sick leave; • know where to find data for calculating these indicators; • be able to independently calculate these indicators.
Abstract	Monitoring of health-related phenomena at workplaces is of utmost importance in the process of making workplaces safer and healthier. There exist several different occupational health indicators. A group of occupational health indicators on sick leave or sickness absence is a group of major indicators which provides information on the health status of the employed people. In this module, four basic indicators are presented indicating occupational health in general namely sick leave percentage, frequency index, severity index, and inactivation index.
Teaching methods	An introductory lecture gives the students insight into the characteristics of basic occupational health indicators on sick leave. The theoretical knowledge is illustrated by a case study. After introductory lectures, students calculate by themselves sick leave percentage, frequency index, severity index, and inactivation index indicators using available data sources.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a lecture room, a computer room; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Calculation of indicators using given data. Only the correct procedure and results of the calculation is considered a positive response.

BASIC OCCUPATIONAL HEALTH INDICATORS ON SICK LEAVE

Marjan Bilban, Lijana Zaletel-Kragelj

Theoretical background

Monitoring of health-related phenomena at workplaces is of utmost importance in the process of making workplaces safer and healthier. Data are needed to determine the magnitude of work-related diseases and injuries, identify workers at greatest risk, and establish prevention priorities, since this kind of morbidity can be prevented. Data are also necessary to measure the effectiveness of prevention activities, and to identify workplace health and safety problems that need further investigation.

Definitions of basic terms

Occupational health

According to World Health Organization (WHO) (1), several definitions of occupational health and safety and occupational health services have been produced by professional bodies, international organizations such as WHO and ILO-International Labour Organization, and different national bodies and authorities. When summarizing these definitions, occupational health is defined as (1):

- protection and promotion of the health of workers by preventing and controlling occupational diseases and accidents and by eliminating occupational factors and conditions hazardous to health and safety at work;
- development and promotion of healthy and safe work, work environments and work organizations;
- enhancement of physical, mental and social well-being of workers and support for the development and maintenance of their working capacity, as well as professional and social development at work;
- enablement of workers to conduct socially and economically productive lives and to contribute positively to sustainable development.

Health indicator

According to A Dictionary of epidemiology (2), health indicator is defined as a variable, susceptible to direct measurement that reflects the state of health of persons in a community. Health indicators are measures which help to compare health status among different populations, or population groups.

Occupational health indicator

An occupational health indicator is a specific measure of a work-related disease or injury, or a factor associated with occupational health, such as workplace exposures, hazards, or interventions, in a specified population (3).

There exist several different occupational health indicators. They could be general, indicating overall occupational health of a population, or specific, indicating occupational health of a specific population group (e.g. males and females, different age groups, etc.).

In this module four basic indicators, presenting occupational health in general, are presented, namely sick leave percentage, frequency index, severity index, and inactivation index (4). All four indicators are indicators of sick leave or sickness absence.

Sick leave

According to Merriam-Webster Online Dictionary, sick leave, in a meaning, used in this module, is an absence from work permitted because of illness (5).

Basic occupational health indicators on sick leave

A group of occupational health indicators on sick leave or sickness absence is a group of major indicators which provide information on the health status of the employed people. Sick leave figures are often used as examples to reveal the need for preventive activities if absence rates are high. At a national level, absence rates are usually examined according to economic sectors to determine what action is necessary. It is also common to consult absence rates at company level in order to determine which departments should be targeted by health promotion activities. The effectiveness of health promotion activities is then often evaluated by the changes in sickness absence rates (6).

Sick leave percentage

Sick leave percentage (SL%) is a percentage of lost days because of a sick leave for one person employed. By sick leave percentage, daily percent of absence is expressed (7). In calculation, calendar days, working days, or effective hours for one employed person could be used.

When sick leave percentage for calendar days is calculated, number of employed persons in the denominator of the equation should be multiplied by the average number of days in a given year, namely 365 (Equation 1a) (4,7):

$$SL\%_{CD} = \frac{N_{lost\,calendar\,days}}{N_{persons\,in\,employment} \times N_{days\,in\,a\,year\,(average)}} \times 100 \qquad \textbf{Equation 1a.}$$

SL = sick leave
CD = calendar days
N = number

When sick leave percentage for working days is calculated, the number of employed persons in the denominator of the equation should be multiplied with the average number of working days, namely 312 or 260 for Slovenia (Equation 1b).

$$SL\%_{WD} = \frac{N_{lost\ working\ days}}{N_{persons\ in\ employment} \times N_{working\ days(average)}} \times 100 \qquad \textbf{Equation 1b.}$$

SL = sick leave
WD = working days
N = number

When sick leave percentage for effective hours is calculated, the number of employed persons in in the denominator of the equation should be multiplied with the number of all effective hours, available for one employed individual (Equation 1c).

$$SL\%_{EH} = \frac{N_{lost\ working\ days}}{N_{persons\ in\ employment} \times N_{all\ effective\ hours}} \times 100 \qquad \textbf{Equation 1c.}$$

SL = sick leave
EH = effective hours (available for one person in employment)
N = number

Frequency index

Frequency index (FI) is the number of employed persons, absent from work because of illness (number of cases) over 100 employed individuals. The frequency index inicates the relative incidence of absence from work because of illness (7). This indicator could be calculated using the following equation (Equation 2) (4, 7):

$$FI = \frac{N_{cases}}{N_{persons\ in\ employment}} \times 100 \qquad \textbf{Equation 2.}$$

FI = frequency index
N = number

Severity index

Severity index (SI) is the number of lost days for one employed individual, absent from work because of illness (number of cases). The severity index indicates the average duration of one absence from work (one person absent from the work because of illness) (7). This indicator could be calculated using the following equation (Equation 3a) (4,7):

$$SI_{CD} = \frac{N_{lost\ calendar\ days}}{N_{cases}} \qquad \textbf{Equation 3a.}$$

SI = severity index
CD = calendar days
N = number

Usually, the severity index is calculated for calendar days, but it could be expressed in terms of working days as well (Equation 3).

$$SI_{WD} = \frac{N_{lost\ working\ days}}{N_{cases}} \qquad \textbf{Equation 3b.}$$

SI = severity index
WD = working days
N = number

Inactivation index

The inactivation index (II) is the number of lost days for one employed individual. The inactivation index expresses the average duration of absence from work (the average number of lost days) for one employed person (7). This indicator could be calculated using the following equation (Equation 4a) (4,7):

$$II_{CD} = \frac{N_{lost\ calendar\ days}}{N_{persons\ in\ employment}} \qquad \textbf{Equation 4a.}$$

II = inactivation index
CD = calendar days

N = number

Usually, the inactivation index is calculated for calendar days, but it could be expressed in terms of working days as well (Equation 4b).

$$II_{WD} = \frac{N_{lost\ working\ days}}{N_{persons\ in\ employment}} \qquad \textbf{Equation 4b.}$$

II = inactivation index
WD = working days
N = number

Case study

Occupational health indicators on sick leave for Slovenia

Data sources for calculating basic occupational health indicators on sick leave

Data on sick leave

In Slovenia, data on sick leave are collected routinely at a national level for all employees in all business sectors.

Routine data for the database are being recorded according to the Healthcare Databases Act (8). The database is entitled "Temporary or permanent sickness absence due to illnesses, injuries, nursing, escort or other causes" (database code in the Healthcare Databases Act: IVZ 3.) (8). They are recorded by medical doctors, authorized for recording these data. Authorization is conceded by the Health Insurance Institute of Slovenia (HIIS). The course of the data is as follows:

- the records, recorded by authorized practitioners are transmitted forward to the nearest Regional Public Health Institute (there are nine regional public health institutes in Slovenia), no later than the 8th day every month;
- data, aggregated for insurance and social medicine analyses at this level, are transmitted forward to the Institute of Public Health of the Republic of Slovenia (IPHRS), no later than the 15th day every month;
- data, aggregated at this level, are transmitted forward to the HIIS, no later than the 20th day every month;
- annual data, aggregated at the regional level are transmitted from regional health institutes forward to the IPHRS no later than February 15th, every year.

Number of employed individuals

In Slovenia, for calculation of indicators on sick leave, the number of employees according to the HIIS database is used. These data are available no later than June 30th, every year.

Very similar data are available in the Statistical Yearbook, issued by the Statistical Office of the Republic of Slovenia (9). These data will be used also in our case study[12].

Reports

All four indicators are routinely calculated by IPHRS. Annual and current three-month reports issued by the IPHRS are published at the web-page of this institution (4).

Basic occupational health indicators on sick leave calculation for 2006

Data, needed for calculation of indicators

Data on the elements for calculation of sick leave percentage, frequency index, severity index, and inactivation index are presented in Table 1[13].

Table 1. Elements for calculation of sick leave percentage, frequency index, severity index, and inactivation index for Slovenia in total for the year 2006

Element	Value
Number of cases*	695,288
Number of lost calendar days*	13,026,763
Number of persons in employment#	824,839

* Data source: National Public Health Institute of Republic of Slovenia (4).
\# Data source: Statistical Yearbook 2007, Statistical Office of the republic of Slovenia (Table 30.17) (9).

[12] Data for all individuals employed except farmers are obtained from the Statistical Register of Employment (SRDAP). Data on farmers are obtained by the Labour Force Survey (LFS). SRDAP covers persons in paid employment who have employment contracts and self-employed persons who have compulsory social insurance (pension, disability and health insurance, parental protection insurance and unemployment insurance). Persons working under copyright contracts, contracts for work/service and citizens of the Republic of Slovenia working in Slovenian enterprises, on construction sites, etc., abroad are not covered.

[13] Owing to slight differences in the number of persons employed between different data sources, values of indicators on sick leave calculated in the case study in this module slightly differ from those published by the National Public Health Institute of the Republic of Slovenia.

Sick leave percentage

For the calculation of SL%, IPHRS is using Equation 1a. In Equation 5 is presented the procedure of calculation of this indicator:

$$SL\%_{CD} = \frac{13{,}026{,}763}{824{,}839 \times 365} \times 100 = 4.33 \qquad \textbf{Equation 5.}$$

Frequency index

For the calculation of FI, IPHRS uses Equation 2. In Equation 6 is presented the procedure of calculation of this indicator:

$$FI = \frac{695{,}288}{824{,}839} \times 100 = 84.29 \qquad \textbf{Equation 6.}$$

Severity index

For the calculation of SI, IPHRS uses Equation 3a. In Equation 7 is presented the procedure of calculation of this indicator:

$$SI_{CD} = \frac{13{,}026{,}763}{695{,}288} = 18.74 \qquad \textbf{Equation 7.}$$

Inactivation index

For the calculation of II, IPHRS employes Equation 4a. In Equation 8 is presented the procedure of calculation of this indicator:

$$II_{CD} = \frac{13{,}026{,}763}{824{,}839} = 15.79 \qquad \textbf{Equation 8.}$$

Exercises

Task 1

Using internet, find sources of data, used in calculation of indicators in Equations 5-8, given in Table 1. Use internet links listed in the Reference list.

Task 2

In Table 2, you will find data for calculation of indicators on sick leave for the year 2006 for twelve statistical regions of Slovenia, separately. Calculate all four indicators on sick leave for each of the twelve statistical regions. Compare results for the twelve specific population groups with the results of the total population and discuss them with your colleagues.

Table 2. Elements for calculation of sick leave percentage, frequency index, severity index, and inactivation index

Statistical region		Number of cases*	Number of lost calendar days*	Number of employed individuals#
1	Pomurska	31,983	608,766	42,253
2	Podravska	102,745	1,899,782	120,513
3	Koroska	21,093	446,621	26,389
4	Savinjska	78,538	1,778,667	104,145
5	Zasavska	7,811	282,156	13,486
6	Spodnjeposavska	21,443	363,319	23,539
7	Jugovzhodna	46,618	894,618	54,044
8	Osrednjeslovenska	196,301	3,836,227	260,344
9	Gorenjska	59,270	1,078,349	71,459
10	Notranjsko-kraska	20,110	307,717	17,557
11	Goriska	59,896	675,091	47,136
12	Obalno-kraska	49,361	852,776	43,975

* Data source: National Public Health Institute of Republic of Slovenia (4).

\# Data source: Statistical Yearbook 2007, Statistical Office of the Republic of Slovenia (Table 30.17) (9).

Task 3

Find appropriate data for calculation of indicators on sick leave for years 2004 and 2005. Compare results for three successive years and discuss them with your colleagues.

References

1. World Health Organization. Global Strategy on Occupational Health for All. The Way to Health at Work. Geneva: World Health Organization; 1995. Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. Accessed: June 4, 2013.
2. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
3. Michigan State University, College of Human Medicine. Putting Data to Work: Occupational Health Indicators; 2005. Available from URL: www.oem.msu.edu/userfiles/file/News/Sv17n1.pdf. Accessed: June 4, 2013.
4. National Public Health Institute of Republic of Slovenia. Sick leave (in Slovene). Available from URL: http://www.ivz.si/Mp.aspx?ni=46&pi=5&_5_id=52&_5_PageIndex=1&_5_groupId=185&_5_newsCategory=&_5_action=ShowNewsFull&pl=46-5.0. Accessed: June 4, 2013.
5. Merriam-Webster Online Dictionary. Available from URL: http://www.merriam-webster.com/dictionary/sick%20leave. Accessed: June 4, 2013.
6. Kreis J, Bodeker W. Indicators for work-related health monitoring in Europe. Essen: BKK Bundesverband; 2004. Available from URL: www.enwhp.org/fileadmin/downloads/memberdocs/580_WORKHEALTH_report_English.pdf. Accessed: June 4, 2013.
7. Bilban M. Statistical indicators on sick leave (in Slovene). In: Terzan M, Pusnik S, Dodic-Fikfak M, Vuga V, Dsuban G, Molan M, Gspan P, Fatur-Videtic A, editors. Occupational medicine for family medicine practitioners. A handbook (in Slovene). Ljubljana: ZZD-Slovensko zdravnisko drustvo; Zdruzenje za medicine dela, prometa in sporta; 2002. p.99-109.
8. Healthcare Databases Act (in Slovene). Ur L RS. 2000;65. Available from URL: http://www.uradni-list.si/1/objava.jsp?urlid=200065&stevilka=2969. Accessed: June 4, 2013.
9. Statistical Office of the republic of Slovenia. Labour force. Statistical yearbook of the Republic of Slovenia. Available from URL: http://www.stat.si/eng/pub_letopis_prva.asp. Accessed: June 4, 2013.

Recommended readings

1. Bilban M. Statistical indicators on sick leave (in Slovene). In: Terzan M, Pusnik S, Dodic-Fikfak M, Vuga V, Dsuban G, Molan M, Gspan P, Fatur-Videtic A, editors. Occupational medicine for family medicine practitioners. A handbook (in Slovene). Ljubljana: ZZD-Slovensko zdravnisko drustvo; Zdruzenje za medicino dela, prometa in sporta; 2002. p.99-109.
2. Kreis J, Bodeker W. Indicators for work-related health monitoring in Europe. Essen: BKK Bundesverband; 2004. Available from URL: www.enwhp.org/fileadmin/downloads/memberdocs/580_WORKHEALTH_report_English.pdf. Accessed: June 4, 2013.

METHODS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**Oral health indicators in Europe**
Module: 2.22	ECTS (suggested): 0.5
Author(s), degrees, institution(s)	Barbara Artnik, DMD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Barbara Artnik Chair of Public Health, Faculty of Medicine, Zaloska 4, SI-1000 Ljubljana, Slovenia E-mail: barbara.artnik@mf.uni-lj.si
Keywords	Dental health surveys, health status indicators, oral health, quality indicators.
Learning objectives	The main two educational objectives of this module are: • to sensitise health professionals for developing an attitude about oral health indicators (OHI) and • to increase the awareness of health professionals about the positive effects of oral health surveys on oral health of a population. After completing this module students should be capable to: • assess the data currently available; • collect additional data according to the new OHI set; • analyse interpret and present the data; and • formulate a policy response to the results.
Abstract	The burden of oral diseases and the needs of populations have been changing rapidly over the past few decades. Therefore, oral health systems are required to adjust to the transition process. In order to meet these challenges effectively, public health care administrators and decision-makers need the tools, capacity and information to assess and monitor health needs, choose intervention strategies, design policy options appropriate to their own circumstances, and improve the performance of the oral health system. The aim of the EGOHIDP – Phase I (European Global Oral Health Indicators Development Project) (2003-2005) started under the European Health Monitoring Programme was to develop a set of indicators for monitoring and describing oral health morbidity and different facets of oral health care systems. As the results a set of 40 indicators in oral public health were identified. The EGOHIDP – Phase II (2006-2008) established methodological criteria for collection of data to implement and promote oral health indicators in an operational way.
Teaching methods	The teaching programme is carried out as a discussion led by moderator. After every activity specific learning objectives will be determined for every participant and until the next workshop their professional tasks should be performed. Their achievements will be reported and discussed with other participants at the next meeting.
Specific recommendations for teachers	Suggested ETCS: 0.5; 2/3 of the work will be done under supervision, 1/3 individual public professionals' work. It is recommended that participants (group of 15 to 20) are all familiar with statistical package SPSS for Windows. A computer room should be provided.
Assessment of students	An attitude test for assessment of attitude changes. The questionnaires applied at the beginning of the first meeting and at the end of the course, or essay, discussing professional impact.

ORAL HEALTH INDICATORS IN EUROPE

Barbara Artnik

Theoretical background

Development of quality indicators

The main dimensions of quality of care include classification of quality into various levels: structure (organisational settings of care), process (health care treatment) and outcome (effects of care), known as indicators, which are divided up as follows (Table 1) (1).

Table 1. Donabedian's dimensions of quality of care

Structure indicators	Process indicators	Outcome indicators
• Resources • Personnel • Equipment • Facilities/Installations • Information systems	• Preventive care • Diagnosis • Therapeutic care • Rehabilitation • Patient information and education	• Health status • Results of care • Patient wellbeing • Patient satisfaction • Efficiency of resource utilisation

An indicator is defined as a variable or parameter, which can measure changes in a phenomenon directly or indirectly, in a valid, objective, sensitive and specific way.

Indicators are an essential component of all phases of health care: policy-making at the health authority level, and treatment and services at the health care provider level. Development of indicators is one of the most significant steps in any quality of care programme, and it is important that those who will primarily be using the indicators in their daily work be directly involved in the process.

Quality indicators are variables whose values indicate the level of quality. Ideally, indicators are related to the final (true) outcome but in some cases intermediate indicators of outcome must be employed. Differing from other methods for evaluating care, the use of true outcome indicators, intermediate outcome indicators and validated structure and process indicators, in the form of quality core data sets, puts the patient at the centre as the key to the successful outcome of care. In the phase of developing quality indicators, it is important that the professional bodies be involved throughout the process, so that there will be agreement on the final selection. Without the support of the professions, the indicators will have little credibility or acceptability. It is easy to see why: if health care providers are to be motivated by their professional pride and satisfaction to improve quality of care, quality indicators will provide the basis of information regarding outcome of care, and providers must see these as relevant, valid and reliable.

In some cases, identifying or defining quality indicators is relatively easy because the literature contains evidence concerning effectiveness of interventions, etc. In other cases, it will be necessary to rely on less validated measures (2).

Overview of oral health indicator sets and oral health goals

Based on this concept, and to enable the evaluation and monitoring of the results of health care, a number of different sets of indicators have been developed; for oral health in 1969, the World Health Organization (WHO) and the Federation Dentaire Internationale (World Dental Federation) (FDI) agreed upon a basic outcome indicator (the number of decayed, missing, filled teeth = DMFT). The first global map with data on DMFT for 12-year-olds showed a high prevalence of caries in industrialized countries and generally low values in the developing countries. A database was established and over a number of years an increasing number of epidemiological studies documented a pattern of change in caries prevalence, i.e. increasing levels of caries in certain developing countries and a decline in caries in many industrialized countries (3,4). Several oral epidemiological studies have been carried out applying WHO methodology and criteria (5).

The caries decline observed in many developed countries (3,4) was the result of a number of public health measures, coupled with changing living conditions, lifestyles and improved self-care practices. In some countries this positive trend could deter action to further improve oral health, or to sustain achievements. It might also lead to the belief that caries problems no longer exist at least in developed countries, resulting in precious resources currently available for caries prevention being diverted to other areas. However, it must be stressed that dental caries, as a disease, is not eradicated but only controlled to a certain degree.

In 1979, the most important goal ever to be formulated for global oral health was announced by WHO. By the year 2000, the global average for dental caries was to be no more than three DMFT at 12 years of age. At the World Health Assembly in 1979, this declaration was unanimously designated as being the overriding priority for WHO. In 1983 oral health was declared to be part of the Strategy for Health for All (resolution WHA36.14) (6). As a part of this goal, the FDI has decided to participate with "goals for oral health in the year 2000". The FDI already has many joint activities with WHO, serving as a link between that organization and the national dental member associations (7). The FDI is thus allowing for the fact that not all recommendations are applicable equally to all countries and populations. Appropriate differentiation is important. In 1989 WHO endorsed the promotion of oral health as an integral part of Health for All by the year 2000 (WHA42.39). In addition, World Health Day in 1994 was dedicated to oral health, which also reflects the importance attached to this issue (6).

In 1981, WHO and the FDI jointly formulated goals for oral health to be achieved by the year 2000, covering the following target age groups: 5–6 years old, 12 years old, 35–44 years old and 65 years old and over, as follows (8):

- 50% of 5-6 year-olds to be free of dental caries;
- The global average to be no more than three DMFT at 12 years of age;
- 85% of the population should have all their teeth at the age of 18 years;
- A 50% reduction in edentulousness among 35–44-year-olds, compared with the 1982 level;
- A 25% reduction in edentulousness at age 65 years and over, compared with the 1982 level;
- A database system for monitoring changes in oral health to be established.

The member countries of the WHO have decided to adopt a global strategy for achieving health for all in the year 2000 (8). Thirty-five European Member States have actually set up national programmes in oral health, and a number of these have reported achieving the target of ≤ 3 DMFT at age 12 (Albania, Bulgaria, Belgium, the Czech Republic, Denmark, Finland, France, Iceland, Ireland, Italy, the Netherlands, Norway, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom) (4,9).

This does not mean that it is possible to standardise concepts of care in all settings. Health care management and concepts vary from country to country and even from region to region within countries and indicators need to be adapted to the local situation. However, the basis remains the same no matter what the local conditions: it is a continuous process which works both through "top down" management and, in reverse, "bottom-up" engagement, which, taken together, can lead to continuous quality of care development (2).

For the new millennium, the WHO Regional Office for Europe specified updated oral health targets (target 8.5) as part of the Health21 policy (10). WHO, FDI and IADR jointly prepared new goals up for oral health in Europe by the year 2020; the indicator used to measure progress - DMFT at age 12 - read: no more than 1.5, and at least 80 % of children aged six years should be free of caries. New oral health goals were urgently needed not only to strengthen dental caries control and prevention activities, but also to address other significant components of the oral disease burden. The objectives and targets have been broadened in order to cover significant indicators related to oral health and care of population groups. The global goals are not intended to be prescriptive but the framework is primarily designed to encourage health policy makers at regional, national and local levels to set standards for oral health in relation to pain, functional disorders, infectious diseases, oro-pharyngeal cancer, oral manifestations of HIV-infection, noma, trauma, cranio-facial anomalies, dental caries, developmental anomalies of teeth, periodontal disease, oral mucosal diseases, salivary gland disorders, tooth loss, health care services, health information systems, and oral health-related quality of life. Such global goals for oral health will assist regions, countries and local health care planners to develop preventive programmes that are targeted at populations and high risk groups, and to further improve the quality of oral health systems.

Priority areas for the European oral health information systems

The burden of oral diseases and the needs of populations have been changing rapidly over the past few decades. Therefore, oral health systems are required to adjust to the transition process. In order to meet these challenges effectively, public health care administrators and decision-makers need the tools, capacity and information to assess and monitor health needs, choose intervention strategies, design policy options appropriate to their own circumstances, and improve the performance of the oral health system. Oral health systems play an important role in establishing optimum oral health by integrating oral health promotion and oral disease prevention into oral health services. Interdisciplinary and intersectional approaches to promotion of oral health have the potential reorienting oral health services towards primary oral health care and that services may better diminish oral disease burdens (11,12).

The global basis for an information system for surveillance of global trends in oral disease and risk factors was introduced by WHO in 1971 but still very few countries have so far prepared a comprehensive national plan for oral health care of the entire population including an information system for monitoring

and evaluation. Therefore it is important to realize that the global goals for oral health in the year 2000 introduced by the WHO and the FDI (8) included a requirement to establish a national monitoring system for oral health. The WHO has developed a simple "Pathfinder" survey methodology (5) for production of baseline data for a national plan for oral health care. This methodology has been successfully applied in a number of countries. As a result the WHO Global Oral Health Data Bank (3) compiles valuable information for monitoring the global epidemiological picture and trends over-time in oral health (13).

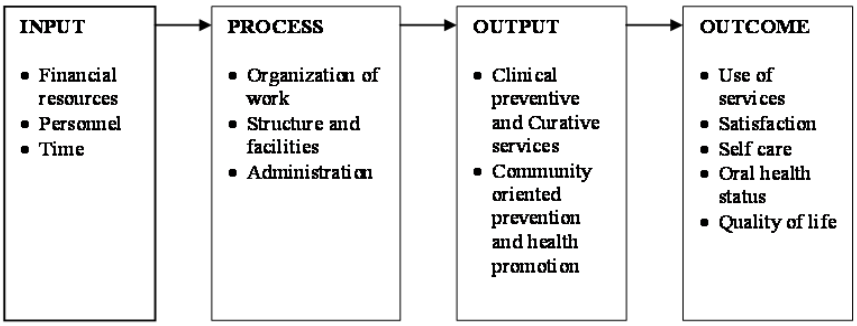
The WHO/FDI goals for oral health by the year 2000 (8) urged Member States to establish oral health information systems, and this remains a challenge for most countries of the world. The WHO Oral Health Programme (14) was prepared to assist countries in their efforts to develop oral health information systems which include data additional to epidemiological indicators.

The information obtainable through a health information system may be usefully categorized into the following interrelated subsystems:

- Epidemiological surveillance,
- Service coverage of the population,
- Service records and reporting,
- Administration and resource management,
- Quality of care provided,
- Oral health programme monitoring and outcome evaluation.

Systematic evaluations of oral health systems are much needed and the WHO Oral Health Programme advocates a comprehensive model whereby input, processes, output and outcomes are measured (Figure 1).

Figure 1. Oral health systems evaluation model. Source: Dr. Poul Erik Petersen, WHO



The WHO Oral Health Programme (14) has initiated integration of the existing database (3) with other WHO health databases and surveillance systems on risk factors. The main surveillance tool is called STEPS (STEPwise approach to surveillance), a simple approach which provides countries with core standardized methods but leaves them flexibility to expand tools by adding information relevant to the local situation (15,16).

The WHO Oral Health Programme (14) provides global health information systems through several activities:

- Revision of the WHO Oral Health Surveys Basic Methods (5), taking new oral disease patterns into account and allowing recording of risk factors to oral health (e.g. dental erosion and consumption of soft drinks);
- Development of procedures for management and analysis of data based on the use of information technology;
- Linking the Global Oral Health Data Bank (3) with the Country/Area Profile Programme information system (4);
- Development of methodologies and approaches for evaluation of effectiveness of community oral health programmes with focus on health promotion and disease prevention. Such evaluation also includes process documentation in order to allow sharing of experiences from programmes.

European Global Oral Health Indicators Development Project 2002 – 2008

Numerous projects have been proposed by different teams from European countries within the framework of the Community Action Programme in the area of health surveillance (17). The European Commission Health Monitoring Programme (18) has as its main objectives to monitor the trends in the European community, to evaluate community programmes and actions and to provide Member States with appropriate health information to make international comparisons and to support their national health policies.

The development of national and international health surveillance systems has resulted in a great number of indicators overwhelming health services personnel in charge of epidemiological surveillance and evaluation of care programmes. The oral health sector is no exception. Within a context of a profusion of health indicators, operating a selection is not an easy task. The need for the necessary integration of the oral health sector within the national and European health information systems is an added challenge, considering that this should be done at all levels of the reference system. A challenge that this European public health project contributes is to meet with practical and decisive recommendations (18).

The EGOHIDP I (European Global Oral Health Indicators Development Project) (SPC 2002472) (12) has been developed in 2002 under the patronage of this Programme (18). The purpose was to establish priorities for a specifically European context in coordination with the existing programme and to make recommendations for improving health system information performance by the establishment of the major indicators of reference.

It was therefore to support the exchange of expectations and experiences among experts of oral health and their audience, policy makers in particular. The terms of reference were also to conduct a systematic review and to outline a process for identifying a set of indicators of reference for oral health that will help national oral health public professionals and services to promote, improve and organize the global oral health promotion, quality of care and surveillance of people in Europe. Overall objectives were listed (12):

- To support European Member States in their efforts to reduce the toll of morbidity, disability related to oral health diseases and especially to strengthen the ability at the local, national, regional levels to measure, compare and determine the effects of oral health services and use of resources on oral health;
- To identify indicators of oral health - problems, determinants and risk factors related to lifestyle - of critical oral health care, its quality of care and of essential health resources; and
- To identify the types of data generation and management problems within the health information system.

Principles for guiding the selection and use of oral health indicators

The major objective of this programme was to contribute to establish a community system for health surveillance (12). It embodied three specific objectives:

- To develop community health indicators through a critical review of existing data and indicators;
- To enable the realisation of a reliable communication system for data and health indicators transfer and sharing;
- To define the necessary methods and instruments for analysis of activities and the production of reports on health status, trends, and policies' impact on health.
- A high priority to identify indicators of reference was to encourage the development of standards for the design and implementation of computerized systems for the management of oral health systems. A goal was to seek a level of agreement sufficient to allow comparability of data that are conceptually equivalent and permit dear delineation of data.

The major principles for guiding the selection and use of oral health indicators focused on:

- The identification of a list of priority oral health problems, populations and high risk group;
- The definition of a table of essential indicators in the following areas: indicators of priority oral health problem, indicators of service delivery, quality of care and indicators of critical health resources;
- The validation of the final long list of oral health indicators;
- A common understanding of term s and criteria for selection of indicators; and
- The recommendation of a short list of essential oral health indicators through a consultation process.

Before even starting to develop a list of existing indicators - a list that should be as comprehensive as possible - the following question should be raised (12) "Which of those indicators are we going to collectively retain on the final list?" As soon as a selection process is engaged, a consensus should be reached on: "What will be the sorting criteria and their hierarchical order?"

This module is restricted to underline the main characteristics for a selection, in relation to the various reference areas: European Community health policies needs, scientific definition, usefulness and feasibility, ethical demand.

European Community Health Policy requirements

According to the European Community Health Indicators Project (12) the oral health indicator set should be:

- Coherent in the sense of conceptual consistency, this implies that a shortlist should nevertheless cover the multidimensional aspect of oral public health surveillance, all areas usually included in the

field of oral public health. This is indeed the fact for the long list already developed which is structured in the four main domains of reference;

- Respond to oral health policy priorities, acknowledging the fact that these will be defined by each Member State and adjusted at local or regional levels;
- Indicators should be scientifically valid, reliable and relevant.

Conceptual consistency

A set of indicators in oral public health, even restricted to a minimal essential list, has a time dimension and should cover the four major following dimensions (12):

- Health status, morbidity and oral function status;
- Determinants (behaviour, life habits);
- Oral health system/promotion, prevention, access to care, quality care and system performance;
- Outcomes and oral health quality of life.

The number of indicators in each area will vary mainly in relation to health policy priorities and to feasibility aspects of data collection and processing.

Methods used

Project members used the following procedure to select the indicators (12): a long list of over 600 possible indicators was drawn up after consultations within the group and with a wide range of relevant European clinical and scientific oral health organisations. Thirty-two group members were then asked to grade the possible indicators in order of importance. Again they were asked to confer widely before making their selections. A statistician then applied the Arrow Theorem (19) to the selections to aggretate the preferences and select 40, which were then discussed by the group.

Descriptions for all the indicators were then written using the structure described later on in this module.

Strengths and limitations of the descriptions of the indicators

The resulting indicators have been selected and described by a process of consensus between a group of decision-makers, clinicians, scientists, administrators and others. All contributors of the group can claim to be “experts” in some areas of oral health but none can claim to be “expert” in all areas. As the evidence base develops in the future and demography and epidemiology change, some of the selected indicators and descriptions will need amendment. However, within the constraints of terms of reference of the European project (18), of time and other resources, the indicators and descriptions do provide a list, which should aid health planners in the future (12).

The issue of health policies

Increasingly European Member States or regions within Members States have formulated health priority areas or targets for health policies. There is a noticeable trend to broaden the spectrum of health objectives moving from simple morbidity measurements, or prevalence of specific diseases to objectives expressed in term s of quality of life improvements, reduction of health inequalities with reference to social policies enabling goals. For example health promotion and prevention tend to focus on specific population group s according to specific life-styles - specifically children or elderly -, goals are formulated for quality of care and access to care, or in terms of social life involvement of entire groups of population such as the aging population. For the oral health sector, this evolution implies a broader concept of the role of oral health professions and their contribution to general health. In addition, special attention should be given to the systematic integration of oral health indicators in any health surveillance system so that trends and changes in life-style and quality of life behaviour in relation to oral health can be monitored effectively. If there is a general move of health strategies towards health promotion and prevention, consideration should nevertheless be given to the fact that the situation varies considerably from country to country. There will be situations for example, where the information priority will be given to the organisation or the reorganisation of the health system for a better quality of care. Clearly health priorities are considerably variable in time and from country to country (12).

Scientific value, reliability and relevance of selected indicators

As short as the list may be, nevertheless, all selected indicators should have the four basic scientific qualities universally accepted. It was proposed to stick to the definitions given by the WHO health statistics programme (20):

- Validity: it is a true expression of the phenomena it is;
- Objectivity: it is able to provide the same result if measured by different people under similar circumstances;
- Sensitivity: it is capable of reflecting changes in the phenomena of interest;
- Specificity: it ret1ects changes in only the specific phenomena of interest.

If the WHO recommendations respond to the necessity of the scientific requirements it is also associated to a deep sense of pragmatism. An indicator that would be qualified “impeccable” scientifically

but too expansive to collect or even impossible to use in a given practical situation would be totally useless. Therefore additional criteria should be considered relating to the actual use of the indicator and to the methodology used to collect the data (12):

- The data required for the indicator are useful for case management or taking action in the community by the staff who originally recorded the data or the service unit from which the data originated;
- It should be feasible to obtain as far as possible through routine service processes or through easily and rapidly executable surveys;
- It should be simple and understandable, measuring one health condition or aspect of the service;
- The indicator and the process of collecting and processing the relevant data are ethical.

Lastly, in the elaboration of the indicators selection process, quantitative principles should be considered as important criteria such as: the frequency of a given health problem, its total costs, its avoidable characteristic (prevention, promotion). This is particularly relevant for indicators of high oral health morbidity and indicators in the field or oral health determinants (12).

A flexible approach to a shortlist of oral health indicators

"STEPwise" approach developed by the WHO (15) is a practical example of a dynamic, multi-dimensional health data collection system, highly adaptable to the objectives and priority information required. In the same spirit, the European Community Health Indicators (ECHI) group proposed the concept of "user-windows" based on the selection of subsets of indicators taken from the comprehensive list of indicators developed (18). The specific user's perspective for selecting user-window could be:

- Specific areas of health policy interest (prevention oriented, services oriented, intersectoral policies),
- Specific thematic entries such as age-groups,
- Specific disease groups with their determinants and costs, etc. This concept offers a more "natural" approach than that of the "core" as the number of possible windows is countless with expansion of information at any level.

Indicators and health objectives

Indicators are markers for health status, system performance and process or available resources. They are usually established to ensure follow-up and evaluation of progression towards health targets formulated by strategic programmes. They should not be confused with public health objectives expressed in terms of disease reduction or public health improvements. These are quantitative measurable achievements reached within a specific time-frame. It should be noted that oral health is broadly integrated within the health sector in the formulation of general targets as well as reflected in the list of proposed indicators. Oral health is considered as a full participative health sector, contributing not only to the promotion of oral health but also as a key actor to the promotion of general health (12).

It should keep in mind that beside their scientific qualities, the selected indicators should (12):

- Respond to the priority needs of the community health strategies, national, local or regional, strategies for disease reduction and health promotion;
- Be practically useful and easy to collect;
- Be part of a highly adaptable information system, adaptable to the variety of needs and resources required for the evolution of scientific and economic contexts.

EGOHIDP I: A set of 40 essential indicators of oral health

As the results of the EGOHODP Phase I (2003 – 2005) which has been supported by DG SANCO, 40 essential indicators of oral health in Europe have been identified and harmonized (12,21,22). They concern problems, determinants and risk factors relating to lifestyle or critical oral health care, its quality of care and of essential health resources. The indicators have been grouped into four categories:

- Children and adolescents (Part A.) (Table 2),
- General population (Part B.) (Table 3),
- Oral health systems (Part C.) (Table 4), and
- Oral health quality of life (Part D.) (Table 5).

Table 2. Twelve indicators for monitoring the oral health of children and adolescents (Part A.)

Group	Indicator
Determinant	A.1. Daily Brushing with Fluoride Toothpaste
	A.2. Preventive Care-Seeking for Pregnant Women
	A.3. Mother's Knowledge of Fluoride Toothpaste for Child Tooth Decay Prevention
	A.4. Fluoridation Exposure Rates
Process	A.5. Preventive Oral Health Programs in Kindergartens
	A.6. Schools with Programs Centred on Daily Brushing with Fluoride Toothpaste
	A.7. Screening Oral Health Programme Coverage
	A.8 Protective Sealants Prevalence
	A.9 Orthodontic Treatment Coverage
Outcome	A.10. Early Childhood Caries
	A.11. Decay Experience in 1st Permanent Molars in Children
	A.12. Dental Fluorosis

Table 3. Eighteen indicators for monitoring the oral health of general population (Part B.)

Group	Indicator
Determinant	B.1. Tobacco Usage Prevalence
	B.2. Daily Intake of Food and Drink
Process	B.3. Geographical Access to Oral Health Care
	B.4. Access to Primary Oral Care Services
	B.5. Dental Contact within the Previous Twelve Months
	B.6. Reason for the Last Visit to the Dentist
	B.7. Reasons for not Visited the Dentist in the Last Two Years
	B.8. Tobacco Use Cessation
	B.9. Untreated Caries Prevalence
	B.10. Removable Denture Prevalence
	B.11. Periodontal Health Assessment
Outcome	B.12. No Obvious Decay Experience
	B.13. Dental Caries Severity
	B.14. Periodontal Diseases Severity
	B.15. Cancer of the Oral Cavity
	B.16. Functional Occlusion
	B.17. Number of Natural Teeth Present
	B.18. Edentulous Prevalence

Table 4. Five indicators for monitoring the oral health systems (Part C.)

Group	Indicator
Determinant	C.1. Cost of Oral Health Services
	C.2. Gross National Product Spent on Oral Health Care Services
Process	C.3. Dentists and Other Oral Care Clinical Providers
	C.4. Dentist Satisfaction with the Quality of Care Given
	C.5. Dentist Satisfaction with the Remuneration Provided

Table 5. Five indicators for monitoring the oral health quality of life (Part D.)

Group	Indicator
Outcome	D.1. Physical Pain due to Oral Health Status
	D.2. Psychological Discomfort due to Oral Health Status
	D.3. Psychological Disability due to Appearance of Teeth or Dentures
	D.4. Social Disability due to Oral Health Status
	D.5. Oral Disadvantage due to Functional Limitation

Indicators for monitoring the oral health of children and adolescents (Table 2) contain a priority list of these indicators which are specific to children and adolescents. It must be appreciated that there are also a range of indicators in "Part B. Indicators for monitoring the oral health of general population" (Table 3) which may also be used to assess oral health in children.

In relation to oral health, "Oral health quality of life" indicators (Table 5) could be defined on the basis of performance indicators (Tables 2–4). For example, elderly people suffering from infected gums and tooth loss will also suffer from diminished quality of life as they will not be able to eat properly and are more likely to be reluctant to socialise as they are embarrassed of their physical appearance.

As described in the WHO Catalogue of Health Indicators (1996) (23), each indicator description includes the following sections:

- Title.
- Rationale. Provides a brief description of the reasons why the indicator has been selected.
- Definition of indicator textually or, in the case of proportions, rate s and ratios, by specifying the numerator and the denominator. The definition should be complete and leave no room for interpretation.
- Definition of important terms, which may have specific meaning in the context of the indicator. Each term in the title of the indicator and its textual definition should be clear to administrative or technical staff not necessarily qualified oral health personnel. Clinical criteria, pathological terms may be defined under this section.
- Common data sources, which could be either routine data collection, special survey or other sources. There may be a need to identify various types of data sources. This section could/should give an indication on how to collect the data (for example as part of community surveys) or where to find already existing information (for example access to databases, review of registers, of patient records etc.).
- Recommended data collection methods which, for some indicators, are specially designed for the needs of the specific indicators.
- Use of the indicator, which is an indication of how the indicator should be used at the facility level, and other levels of the health system. For example: to identify high-risk groups for implementation of preventive programme.
- Recommended formats of presentation.
- References providing primary sources of additional information about this and possibly other related indicators.

The report on selecting essential oral health indicators in Europe issued from the EGOHIDP I (12,21,22), supported by the Health and Consumer Protection Directorate-General, European Commission, called upon policy-makers, community leaders, private industry, health professionals, the media, and the public at large to affirm that oral health is essential to general health and well-being and to take action. Particular attention is given to the negative implications to health of changing diet, nutrition and unhealthy lifestyles related to tobacco and excessive use of alcohol. Moreover, priority should be given to the problems emerging within deprived communities or disadvantaged populations in Europe.

EGOHIDP II: Methodological criteria

EGOHIDP Phase II (2006 – 2008), according to existing morbidity projects DG SANCO (ECHI, EUROHIS, ISARE, etc.), established methodological criteria for the collection of data to implement and promote oral health indicators in an operational way in order to be able to support and achieve the overall objectives (24,25). The EGOHIDP I (12,21,22) has enabled a feasibility study (EGOHIDP II) (24,25), which is an essential part of an overall project, since it will allow Member States to evaluate their capability to use these indicators.

With the support from the European Commission, the expert contribution of the ministries of health, universities, regional and national dental associations, health professionals in the European member states, the EGOHID will facilitate service specifications across area health services with a view to maintaining and improving performance and with the enhancement of the capacity to analyze the social, economic, behavioural and political determinants with particular reference to poor and disadvantaged populations.

Since indicators are an essential requirement for comparisons to be made over time not only between regions and care units but also at national level, these comparisons can be used as a basis in development and quality work at all levels of dental care and dental services. Although the prerequisites for monitoring the quality of care in Europe are good, despite major disparities between Members States, further development and promotion of models and methods for performance assessment is needed in order to be able to deliver policy-relevant information to the nation's health policy makers (26).

Exercises

For the purposes of this training programme four tasks will be executed (according to the learning objectives). The whole programme will be carried out as a discussion led by moderator. After every task specific learning objectives will be determined for every participant and until the next meeting their professional tasks should be performed. Their achievements will be reported (within 10 minutes) and discussed with other participants at the next meeting.

Task 1

Stimulating introduction at the first meeting will be led by moderator: key words will be used as a target to sensitise the participants that good oral health is essential for good general health. Discussion: The assessment process of the availability of data. Task 1 they have to achieve until the Meeting 2:

- To inventory the data that are already being collected and that can be used to assess the oral health status in different population groups;
- To assess the informative value of these data;
- To make provisions for generating new data.

Task 2

At the second meeting the reports should be presented by every participant. Discussion: Oral health indicators. The results of the first workshop will determine whether additional indicators need to be collected.

Task 2 they have to achieve until the Meeting 3 (if necessary):

- To add indicators to existing data sources;

Task 3

At the third meeting the reports should be presented by every participant. Methodological guidelines should be discussed and refined. It has to be decided:

- Which oral health indicators will be used;
- The choice of an adequate level of analysis and the application of multilevel analysis.

Task 3 they have to achieve until the Meeting 4:

- To analyse differences in oral health;
- To interpreted the results carefully;
- To prepare the results for clear and understandable presentation.

Task 4

At the fourth meeting the results have to be presented clearly and understandably (e.g. to use graphical displays) by every participant. The discussion: Formulating public health policy priorities to the results:

- To what extend has the state identified oral health as an important part of general health until now;
- What are the objectives for any interventions;
- Who are the main groups with a concern for poor oral health;
- What are their interests, priorities, and commitments;
- What is the context within which interventions need to be considered;
- etc.

The formulated document should assure that public health policy satisfies identified needs and finally it should be submitted to policy-makers.

Follow up workshops on health policy development should be performed every six months.

References

1. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743–8.
2. WHO Regional Office for Europe. Joint ECE-WHO Meeting on Health Statistics, 14 – 16 Oct 1998. Available from: www.unece.org/stats/documents/ces/ac.36/1998/16.e.doc (Accessed: June 5, 2013).
3. World Health Organization. Global Oral Health Data Bank. Geneva: World Health Organization, 2002.
4. World Health Organization. WHO Oral Health Country/Area Profile. Available from the Internet on the Internet from: http://www.who.int/oral_health/databases/malmo/en/ (Accessed: June 5, 2013).
5. World Health Organization. Oral health surveys: Basic methods. 4th ed. Geneva: World Health Organization, 1997.
6. Petersen PE. The World Oral Health Report 2003. Continuous improvement of oral health in the 21st century - the approach of the WHO Global Oral Health Programme. Geneva: World health Organization, 2003:38.
7. Aggeryd T. Goals for oral health in the year 2000: cooperation between WHO, FDI and the national dental associations. Int Dent J 1983;33(1):55–9.
8. World Health Organization/Federation Dentaire Internationale. Global goals for oral health in the year 2000. Int Dent J 1982;32:74–7.

9. World Health Organization. Health for All: list of indicators. Copenhagen: World Health Organization, Regional Office for Europe, 1993.
10. World Health Organization, Regional Office for Europe. Health21. The health for all policy framework for the WHO European Region. European Health for All Series No. 6. Copenhagen: World Health Organization, Regional Office for Europe, 1999:224.
11. World Health Organization. Information systems. Oral health information systems, evidence for oral health policy and formulation of goals. Available from: http://www.who.int/oral_health/action/information/surveillance/en/index1.html (Accessed: June 5, 2013).
12. Bourgeois DM. Health surveillance in Europe. European Global Oral Health Indicators Development Project. SPC 2002472. Final report 2003 – 2005. Lyon: Community Action Programe on Health Monitoring, European Commision, Health and Consumer Protection Directorate-General. Available from:

    http://ec.europa.eu/health/ph_projects/2002/monitoring/fp_monitoring_2002_frep_03b_en.pdf (Accessed: June 5, 2013).
13. Tala H. Information systems for oral health. Int Dent J 1987;37(4):215–7.
14. World Health Organization. WHO Oral Health Programme: Available from: http://www.who.int/oral_health/en/index.html (Accessed: June 5, 2013).
15. Bonita R, de Courten M, Dwyer T, Jamrozik K, Winkelmann R. Surveillance of Risk Factors for noncommunicable diseases: The WHO STEPwise approach. Summary. Geneva, World Health Organization, 2001.
16. Bonita R, Strong K. The SURF Report I. Surveillance of risk factors related to noncommunicable diseases: current status of global data. Geneva: World Health Organization, 2003.
17. European Commission. Decision No 1786/2002/CE of the European Parliament and the Council of 23 September 2002 adopting a programme of Community action in the field of public health (2003 – 2008). Official Journal of the European Communities No. L 271, 9 October 2002.
18. ECHI. Design for a set of European Community Health Indicators. Final report of the ECHI project under the EU Health. Monitoring Programme. Bilthoven: RIVM, February, 2001.
19. Arrow KJ. Social Choice and Individual Values. 2nd ed., 8th printing. New Haven, London: Yale University Press, 1976, cop. 1951.
20. World Health Organization. Development of health programme evaluation: Report by the Director-General. Geneva: World Health Organization, 1978 (document A31/10).
21. European Commission. Bourgeois DM, Llodra JC, Norblad A, Pitts NB, eds. Health surveillance in Europe. A selection of oral health indicators; recommended by the European Global Oral Health Indicators Development Project. 2005 Catalogue. Lyon: University of Lyon, 2005.
22. Bourgeois DM, Llodra JC, Nordblad A, Pitts NB. Report of the EGOHID I Project. Selecting a coherent set of indicators for monitoring and evaluating oral health in Europe: criteria, methods and results from the EGOHID I project. Community Dent Health 2008;25(1):4–10.
23. World Health Organization. Catalogue of Health Indicators. A selection of important health indicators recommended by WHO Programmes. WHO/HST/SCI, Geneva: World Health Organization, 1996.
24. European Commission. Bourgeois DM, Ottolenghi L, eds. Health surveillance in Europe 2008. Oral health interviews and clinical surveys: overviews. Milan: Quintessence International, 2008.
25. European Commission. Bourgeois DM, Christensen LB, Ottolenghi L, Llodra JC, Pitts NB, Senakola E, eds. Health surveillance in Europe. Oral health interviews and clinical surveys: guidelines. Lyon: University of Lyon I, 2008.
26. European Commission. EGOHID Phase II – Summary. Available from: http://www.egohid.eu/html/Phase2_summary.htm (Accessed: June 9, 2013).

Recommended readings

1. McKee M, Ryan J. Monitoring Health in Europe: opportunities, challenges, and progress. Eur J Public Health 2003;13(3 Suppl):1–4.
2. Ochoa A, Imbert F, Ledesert B, Pitard A, Grimaud O. Health indicators in the European regions. Eur J Public Health 2003;13:118–9.

3. World Health Organization. WHO HFA indicators for the new health policy in Europe. Report on a WHO Expert Group Meeting, The Hague, Netherlands, 2-3 March 2000. Copenhagen: World Health Organization, Regional Office for Europe, 2000.
4. World Health Organization, European Commission. International Compendium of Health Indicators (ICHI), version 1.1. Copenhagen: World Health Organization, Regional Office for Europe, 1999.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Outbreak investigation**
Module: 2.23	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Maja Socan, MD, PhD, Assistant Professor National Institute of Public Health, Ljubljana, Slovenia
Address for correspondence	Maja Socan National Institute of Public Health Trubarjeva 2, 1000 Ljubljana, Slovenia Tel: 0386 1 2441 522; e-mail: maja.socan@ivz-rs.si
Keywords	Communicable diseases, outbreak investigation.
Learning objectives	After completing this module students should: • recognise the ongoing outbreak and the urgency of immediate public health action to prevent the spread of communicable diseases; • define an outbreak and describe the steps in an outbreak investigation; • understand the importance of a case definition in the context of an outbreak investigation, construct and interpret an epidemic curve and calculate attack rates; • characterise an outbreak by time, place and person; • be capable to perform field investigation and analysis of data collected; • increase knowledge in measures for the control of communicable disease outbreak and the need of follow-up of control measures.
Abstract	An outbreak occurs when there are significantly more cases of the same communicable diseases than past experience would have predicted for the time, place and among the population under observation. The phases in outbreak investigation are: defining the problem, appraising the existent data – including case identification, clinical observation, tabulation of data collected, identification of microorganism causing communicable disease, formulating the hypothesis, testing the hypothesis, drawing the conclusions and preparation of control measures. The module is describing principals of outbreak investigation with implementation and follow-up of control measures.
Teaching methods	Teaching methods include introductory lectures on phases of outbreak investigation. An open discussion follows the lectures to clarify some points and give an additional explanation to students if needed. Students have to prepare a detailed work plan of outbreak investigation with operations at central level, field investigation, analysis of investigation data and measures for the control of an outbreak. The work plan is at first discussed in a small group, then presented to other students. The discussion should follow with emphasis on how practical problems should be managed and solved. Students prepare a complete written report on outbreak investigation which is read and corrected by the teacher. Students have to find the published reports on similar outbreaks and compare the solutions proposed for their containment.
Specific recommendations for teachers	• Work under teacher supervision/individual students' work proportion: 30%/70%; • equipment: computers (one computer for 2-3 students), LCD projection equipment, internet connection, access to the bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire; Structured essay; Seminar paper; Case problem presentations.

OUTBREAK INVESTIGATION

Maja Socan

Theoretical background

An outbreak appears suddenly and unexpectedly. The public health measures should be implemented as soon as possible to limit the number of new cases (1,2). At the beginning, the population affected, the causative agent, the mode of transmission is unknown in most of the outbreaks. Ongoing outbreak demands a quick response by public health experts and authorities (3,4).

The possibility of increased number of communicable disease cases demands a thorough outbreak investigation which is the responsibility of public health (5). The success of the investigation will depend on the methodical organization of operations both at central level and in the field.

The crucial steps in epidemiological investigation are (1):

- defining the problem;
- appraising the existent data – including case identification, clinical observation, tabulation and spot maps, if necessary, and identification of the microbe;
- formulating the hypothesis;
- testing the hypothesis;
- drawing conclusions and prepare control measures.

Operations at central level

As soon as the initial information on an outbreak reaches the public health institution, the epidemiologist must verify the information. If the existence of an outbreak is confirmed (or even pending confirmation) the epidemiologist must analyze the situation and initiate the decision-making processes at the central level (6).

The epidemiological team(s) for on-the-spot investigations should be set with no delay. The team of the epidemiologists must have very clear instructions (7).

Checking of initial information on an outbreak

The information on an outbreak may come initially from different sources (1):

1. routine or emergency reports from medical facilities;
2. routine epidemiological surveillance;
3. the early warning system;
4. from other sources, e.g. veterinary services, laboratories, or, as frequently happens, from rumours disseminated rapidly by the media or by people who sensed that a problem existed (8).

The epidemiologist checks carefully the validity of the information on the outbreak. Comparison of the information obtained from a number of different sources will indicate whether the initial reports are reliable. The situation may also require a rapid site visit by a competent person who should have both clinical and epidemiological experience of the suspected disease. The epidemiologist has to be aware of different agents involved in the outbreak and collect appropriate laboratory specimens to confirm the tentative clinical diagnosis. The broad differential diagnosis must be kept in mind before a preliminary conclusion is reached (9).

Preliminary analysis of the situation

As a first step, it is necessary:

- to establish a clear definition of the disease (case definition);
- to formulate the initial hypotheses regarding the nature of the agent and the cause of the outbreak;
- to formulate the objectives and strategy of the outbreak investigation.

Case definition

The case definition is the crucial for outbreak investigation as it will serve as a guide in case finding. The case definition should be precise but at the same time not too exclusive. The provisional case definition will be based on the examination of the earliest cases seen and will be changed as soon as more precise clinical and epidemiological data become available (6). A final case definition should contain the following:

- the name of the disease (syndrome);
- the most frequent and the occasional signs and symptoms in both mild and severe cases;
- the circumstances associated with the occurrence of cases (e.g. time, place, habit, contact, etc.);
- a suggestive or confirmatory laboratory test;
- the criteria for “confirmed”, “'suspected”, and “probable” case, and index case, “primary” or “secondary” position in the transmission chain.

Since not all cases will be investigated with the same degree of thoroughness, a grading system is necessary to indicate the certainty with which the diagnosis is made. The criteria for confirmed, suspected and probable cases are defined (6).

The following is an example of a case definition in an outbreak of a suspected food-borne disease to which the name dysentery-like syndrome was given. Initial definition (Example 1):

Example 1. A person having diarrhea and one or more of the following signs and symptoms: fever, nausea, vomiting, abdominal cramps, tenesmus, blood in the stools.

The final definition is made in the end of the investigation and can be graded:

- by severity: a severe case is one with fever above 38ºC and bloody diarrhea, with or without other signs and symptoms as indicated above, a mild case is one fitting the initial definition, but without high fever and bloody diarrhea;
- by level of certainty: a confirmed case is one from which a strain of Shigella has been isolated and identified, with or without clinical signs or symptoms, a suspected case is one where no agent has been isolated but where the faecal exudate is rich in polymorphonuclear leukocytes or macrophages, a probable case is one with a compatible clinical picture, but without positive laboratory findings;
- by epidemiological associations: a primary case is someone who ate a certain food with other primary cases and became ill.

Formulation of hypotheses

Initial hypotheses should be formulated as to:

- the nature of the disease;
- the origin of the outbreak;
- the mode of transmission.

Such hypotheses are necessarily based on incomplete information but also essential as a guide to further investigations (7). The hypotheses are subjected to modification, refinement, or total change as the investigation proceeds. The approach by syndrome to the etiological diagnosis of disease causing outbreak, facilitates a comprehensive review of the different microbes that should be considered in the differential diagnosis.

Equipment and logistic support of outbreak investigation

Rapidity and efficacy are essential to successful outbreak investigation. The contingency plan should be prepared in advance to help in making available in the shortest possible time all the equipment and logistic support needed by field teams. The types of equipment that may be required will depend on local conditions (10).

Maximum protection, including respiratory, may be necessary in examining highly contagious patients, and in high-risk operations, such as post-mortems or the processing of dangerous laboratory specimens. Gowns, gloves and surgical masks can be effective in preventing disease transmission (10).

Laboratory support

Laboratories must be identified within the country, which are capable of diagnostic procedures (1). In decision process to which laboratory, or laboratories, specimens should be sent, the following should bet taken into consideration:

- the nature of the suspected microbe;
- the level of expertise required and available;
- the types of protective equipment available;
- the facilities for the shipment of specimens;
- the delays expected in receiving results.

Organization of field investigation

The organization of field investigation includes selection of investigation team, equipment and logistic support and depends on the time required to find and examine the cases, the time required to collect laboratory specimens, the extend of the outbreak and the time required for completing emergency control measures (2).

Outbreak investigation team should be given precise instructions covering their activities, including safety precautions, methods of case finding, contact tracing, special investigations, collection and shipment of laboratory specimens (7).

Case finding

During outbreak investigation, the procedures used are aimed to identify as many cases as possible. The use of a carefully designed questionnaire form is necessary to ensure accurate and rapid investigation (6). Records of cases should be based on precise definitions and the findings systematically validated.

Case description and recording

A standard case investigation form is needed to ensure that complete information is obtained for each case. The type of information needed is the same for all diseases outbreaks, but specific details must be adapted to the individual disease and to the unique circumstances of each outbreak situation and each location (11). Despite the pressure of time, great attention must be given to patient identity, the serial numbers allocated to cases, forms and specimens, and the labeling of specimens. Important data are too often rendered meaningless or misleading because of errors (12).

In some cases, data should be collected from hospitals and community health centres. A retrospective survey should be conducted using records of inpatients, outpatients and laboratory results going back over the previous three months or so (13). Special attention should be given to cases that might have been misdiagnosed. In any case, an increase in the number of consultations or admissions should arouse suspicion.

Community survey is indicated not only to discover suspected cases but also to investigate the epidemiological factors that may have contributed to the spread of the causative agent in the community. The community may be a city, district, village or camp, depending on the circumstances (14). A case-finding strategy should be developed in the community concerned, using one or more of the following procedures (8):

- visit to local health facilities;
- interviews with doctors, pharmacists, nurses, veterinarians, and veterinary health personnel etc.;
- interviews with a random statistical sample of the population or with the population at points where people gather together;
- visit to hospitals known contacts of inpatients and outpatients who reported to hospitals and health centres;
- systematic house-to-house visits (extensive or to a random statistical sample).

Search for source of infection and contact tracing

The primary of searching for the source of infection, either of an individual case or of the entire group of cases, is to eliminate, terminate, or isolate the source so that similar circumstances do not occur again or are less likely to do so in the future.

The methods used for identifying sources and tracing will differ according to whether an individual case or an outbreak is being investigated, whether the relevant infection is transmitted from person to person or by common source origin (3). However, the steps to be taken and the order in which they are taken will remain the same:

- identify the date (or time) of disease onset;
- ascertain the range of incubation periods for the disease in question,
- look for a source of infection in the time interval between the maximum and minimum incubation periods.

Incubation periods vary from as little as a few hours (e.g., salmonellosis), to days (e.g. influenza), weeks (e.g. hepatitis A), and even months (e.g. rabies) (6).

Outbreak can be:

- with continuous person-to-person transmission;
- with a common source of infection.

When communicable disease causing outbreak may be transmitted by some other mechanism, a search should be made at the times indicated by the incubation period for the presence of vectors, reservoir animals, environmental contamination, or whatever the appropriate source of infection is (5).

If the analysis provides grounds for suspecting that a common source was involved, the search for that source is basically the same as in person-to-person transmission. The period of time during which the common exposure might have taken place is determined by the range of the incubation period and this in turn defines the period to be covered by the search (6). The common source may have been an individual and the mechanism of transmission person-to-person (9). Information on exposure history is obtained from the index case. As previously mentioned, this section of the case investigation form must be specially designed for a particular disease outbreak.

Continued transmission

The purpose of prospective or potential case contact tracing is the identification of new cases that may already have occurred or may still result from contact with the source of infection. Whether the mechanism of spread is person-to-person, via vectors, or contaminated food or the environment, the investigation should be based on the infective microbe (11). As the infective period may be brief (e.g influenza), a few weeks (e.g., hepatitis A), a year or more (e.g. malaria), or lifelong (in chronic carriers of hepatitis B) the maximum duration of the infective period defines the time during which a search for new cases should be conducted (6).

It is important to recognize that not only does the duration of the period of communicability vary, but also that this period may or may not begin before the appearance of signs and symptoms, and may or may not continue for a variable time after they have appeared (11).

Forward (prospective) tracing of contacts has two main purposes:

- to identify continuing chains of infection and/or contamination, in order to interrupt them;
- to locate new cases, so that they can be treated and further spread of infection discontinued.

It is particularly important if the patients have travelled outside his/hers own community, and may therefore have introduced infection some distance away from his home.

Special investigations for common-source infections

Whenever there is some evidence of a common source of infection, special investigations must be carried out. Such sources may include arthropods, vertebrate animals, food, beverages and the environment.

Food-borne diseases

A food-borne disease will obviously be suspected when a number of persons who have eaten a meal together fall ill (15-19). Finding the infected dish responsible is more difficult and all those who eat the meal should be classified into subgroups according to the dishes that they consumed (16). Tracing the source of infection is even more difficult if the incriminated food has been eaten intermittently in different places, or if the contaminated product is mixed with different kinds of food or beverages (17). Food contamination may originate from infected animals, food handlers, flies and the environment.

Investigations may have to be carried out on:

- the conditions under which the food concerned is grown, produced and consumed;
- the handling and storage of foodstuffs, with particular reference to conditions known to be potentially hazardous;
- the sanitary condition of restaurants, hotel etc., and their use by the community;
- the possibility of cross-contamination from raw food to cooked food, the packages, bags, or containers in which the food has been transported, and the cooking utensils and working surfaces associated with its preparation;
- the health status and hygienic practices of food handlers.

The assistance of the veterinary services may also be needed here (17).

Arthropod-borne diseases

Numerous blood-sucking insects are able to transmit diseases from man or from animal reservoirs to a exposed person: ticks, mosquitoes, fleas, body lice, mites, sand flies etc. They are most often specialized in the diseases that they can transmit and this knowledge, plus their abundance at the time of the outbreak, can lead to a particular insect being incriminated as the mode of transmission (20-26). The possible role of insects is best confirmed by a specialists entomologists-epidemiologist, who will also know how best to collect specimens for laboratory examination.

Zoonoses

Zoonoses may be transmitted from vertebrate animals to man both by arthropod bites and by contamination of food and the environment. Direct contact with sick domestic or wild animals or healthy carriers, such as horses, sheep, cattle, goats, pigs, dogs, cats, poultry, monkeys, rodents, and birds, may cause the epidemic diseases (27-29). Proof that infection in an animal species is related to the human disease may require the assistance of the veterinary services.

Originating in the environment

Common-source-infection may originate in the environment if water, soil or air is contaminated; the source of such contamination may be man or animals (30-33). In addition, certain agents of mycotic diseases may be present in soil. Several modes of transmission may be involved:

- water (water-borne diseases): transmission via drinking water, beverages, contaminated food, or from bathing in recreational waters;
- soil: transmission by direct contact, or contact of dust with mucous membranes (respiratory tract, eyes);
- air: inhalation of drops or aerosols.

Proof of causation should be carefully established, and this may be very difficult. It may be necessary to investigate water-treatment systems, sewage systems, possible contamination of water and soil by human and animal faeces, use of soil as fertilizer, water recycling, and air conditioning systems (33).

Collection and shipment of laboratory specimens

As has been noted in previous sections, laboratory support is essential in clinical and epidemiological investigations. The value of the results obtained will depend on:

- correct sampling of appropriate specimens;
- correct storage, packaging, and shipment;

- appropriate formulation of requests for laboratory examinations;
- the speed with which the laboratory responds to such requests.

If possible, therefore, the field team should include a microbiologist, or seek the advice of one whenever necessary (34).

Analysis of investigation data

After the field outbreak teams have finished their investigations, clinical, epidemiological and laboratory data is compiled and analyzed as soon as they become available. The data collected are used to arrive at a probable clinical diagnosis of the disease, define the epidemiological characteristics of the outbreak, confirm the identity of the causative agent, and identify the appropriate control methods (35).

Clinical data, i.e., the signs and symptoms recorded for each patient, are tabulated, the more precise picture of the disease thus obtained enabling the provisional case definition to be revised and providing a clinical approach to an etiological diagnosis. A disease is generally described either in terms of the relative frequency of the various signs and symptoms that have been observed, or by drawing the way that these frequencies change during the course of the disease. If the number of person examined is large enough, the data for suspected, presumptive, and confirmed cases, or for mild and severe forms of the disease, can be tabulated separately (6).

Without microbiological confirmation, a syndrome can be identified (e.g. febrile rash, hemorrhagic fever, febrile lymphadenopathy, febrile neurologic disease etc.) and possible causative agents listed.

Epidemiological data are required to gain knowledge, who is affected, what is the size of the outbreak and to formulate a hypothesis as to the causative agent, the mode of transmission, and the probable progress of the outbreak (36,37).

Data on the numbers of persons affected are relatively meaningless for the purposes of epidemiological analyzes without knowing the number of persons exposed or at risk. Rates must be therefore calculated but only if numerators and denominators are known (38).

Attack rates and case-fatality rates are calculated by personal characteristics e.g. age, sex and occupation. Spot map and distribution of cases in time are equally important indicators (39,40). The distribution of cases in time is best shown by a graph (histogram).

Contact tracing usually points out to one of the transmission patterns (person-to-person, etc.). A hypothesis of causation is deduced and confirmed by statistical analysis.

General measures for the control of outbreaks

An outbreak of communicable disease may be controlled by:

- eliminating or reducing the source of infection;
- interrupting transmission;
- protecting persons at risk.

It may take some time before the exact nature of the causative agent is known and this will delay the application of specific control measures, such as immunization of persons at risk or the treatment of carriers. In an emergency, therefore, the first step must be to try to interrupt transmission, since the epidemiological investigations will quickly provide some indication of the possible mode of transmission involved (41-43). This may be:

- person-to-person transmission, whether direct or indirect;
- common source of infection;
- combination of both.

General measures to be taken in various types of outbreak are described. In emergency conditions, control measures may require a degree of improvisation whenever the necessary equipment is not immediately available; this is not difficult when the principles to be followed are well understood (43).

Measures in outbreaks of diseases with person-to-person transmission

Measures may be necessary in respect of patients, their contacts and the community.

Patients

The health personnel participating in medical care, specimen collection, laboratory examination, post-mortems, and field operations during outbreak investigations will all require protection. Immune personnel (after immunization or natural infection) should be employed if possible. However, when the agent is unknown or if there is no vaccine, general precautions are indicated which must be adapted to the degree of contagiousness of the disease; they should be reliable but not excessive, so as to avoid waste of time and money.

The most effective general precaution is careful hand-washing after any contact with a patient, or with a suspected case. Protective measures may be divided into four categories, depending to the degree of communicability of the disease and its mode of transmission, as determined by the outbreak investigation (41,42).

Duration of precautions or isolation has to be determined. The infective or contagious period is known for most communicable diseases. When the agent is unknown, the period of contagiousness can be determined from the data of infective contacts collected during the outbreak investigation, which may fit one of a number of different patterns (36,37).

Disinfection

Safe disposal of excreta infectious material (vomit, urine, secretions, discharges, dressings and bedding) is recommended, and may be mandatory, depending on the mode of transmission of the disease; this may be achieved by using disinfectants or by incineration (14). If contaminated material is to be transported, the double-bagging procedure must be used.

Contacts

Persons who are in contact with a patient with communicable disease during the contagious period may be at risk of becoming infected and therefore of becoming in their turn a source of infection (9,10). However, the magnitude of this risk is not the same for all diseases and for all persons, and must therefore be assessed and preventive measures adapted accordingly.

Assessment of the risk of communicable disease

The following factors influence the risk of infection: the time of contact, and in particular whether it falls within the period of contagiousness, the degree of contagiousness of the disease, the closeness of contact and the routes of transmission to which the person may have been exposed and the specific and non-specific immunity of the person concerned (11).

During control operations, the time of contact and the closeness of contact are the essential factors in determining the measures to be taken. Two types of contact may be distinguished (6):

1. A close contact.
   A close contact is a person who has had occasional face-to-face contact, has given personal care without protection measures, or has shared the same meal or room during the period of communicability, or handled the patient's belongings (if indirect transmission is involved);
2. A possible contact.
   A possible contact is a person who may have been exposed either:
   - at some distance away from a highly contagious case during the period of communicability in circumstances not satisfying the above criteria, e.g, in public transport, in the next bed in a hospital, or in the same workplace; or
   - thought close contact with a patient outside the period of communicability, particularly if there is some doubt about its duration.

Quarantine

This is used to restrict the contacts of a well person who has been exposed to a patient with a communicable disease during the communicability period. Quarantine must be adapted to the risk to which the person concerned was exposed and the risk that he represents for the community (1). The restrictions imposed should not be excessive from either the humanitarian or economic point of view. A large number of contacts may have to be dealt with in a few days. They should therefore be divided into "possible contact" and "close contact" groups, which should be dealt with separately. Each group should be divided into cohorts depending on the expected time of onset of the disease concerned; this will be determined by the range of incubation periods following the infective contact. When there are numerous contacts should be separated physically so as to avoid introducing new suspects into a group that has already completed part of the quarantine period and who should then be obliged to begin the whole period again.

The protection of patients and the isolation of their contacts in quarantine will considerably decrease the risk for the community. However, as it may not be possible to identify all patients and contacts, other methods also have to be considered.

Immunization

Emergency mass immunization is possible for a limited number of diseases (1,6). But there will inevitably be some delay before a large enough part of the population is protected by the vaccine; other methods may therefore be necessary during the interim period.

Chemoprophylaxis

Chemoprophylaxis may sometimes be used during outbreaks to protect persons who have been in contact with source of infection or with infected person (1,6).

Public health measures

Public health measures are as follows (2):

1. Restrictions on mass gatherings:
   Such restrictions may be indicated including the closure of schools and even of public places, but their effectiveness is generally limited.

2. Restrictions on travel:
   This may involve the establishment of a cordon sanitaire in order to isolate the epidemic focus or to prevent the entry of infectious persons into a country. There is, however, more justification for a cordon sanitaire when immunization is possible and the aim is to make sure that unimmunized persons do not travel and thereby carry the disease to other places. Before a cordon sanitaire around an epidemic focus can be established, it is first necessary to define the boundaries of both the infected and the receptive areas. This is expensive, and requires close cooperation between the health services, the police, and the army, without which the measure may be ineffective. Furthermore, considerable economic loss and inconvenience may be caused to individuals.
3. Strengthening of epidemiological surveillance:
   This has proved to be both more efficient and less expensive than the cordon sanitaire. Case finding, contact tracing, and prevention of transmission should all be strengthened in any group in which suspected cases have appeared (8).
4. Community participation:
   Keeping the community informed will reduce the risk of panic. If the community can be included to participate in the control measures, this will contribute considerably to their effectiveness.

Control of outbreaks caused by a common source of infection

Whenever an outbreak is caused by a common source of infection - whether by arthopods, rodents, direct contact with vertebrate animals, food, water, air, soil or a combination of any of these - control methods should be based on source reduction and interruption of transmission. The assistance of a specialist entomologist, veterinarian or sanitary engineer may be required.

Mosquito-borne diseases

Mosquitoes are capable of transmitting diseases to man belong to several species and their control raises technical problems that require the assistance of a specialized team (20). They constitute the most important group of insects vectors, transmitting malaria and a number of arboviruses, including those causing outbreaks of yellow fever, dengue and dengue haemorrhagic fever, Japanese encephalitis, New Word equine encephalitides, and several dengue-like fevers (21,22). Only the females bite humans. They lay their eggs in impounded water, selected according to the preference of the species. Mosquito control requires planning of strategy, logistic, and field operations. It should be noted that a patient with a mosquito-borne disease, e.g., dengue or yellow fever in Aedes aegypti infested areas or malaria in Anopheles-infested areas, should not be moved into an area where such mosquitoes are present; such movement may be subject to local health regulations.

Rodent-borne diseases

Rodents may be reservoirs of a number of epidemic diseases, including leptospirosis, plague, tularaemia, yersiniosis, lymphocytic choriomeningitis, Lassa fever, Junin and Machupo haemorrhagic fever with renal syndrome. Certain rodent-borne diseases may be passed from rodents to human by “'direct” transmission, others thought arthropod vectors. "Direct" transmission occurs as a result of contamination of food and water by rodent urine and can thus also be regarded as indirect (44,45). The results of the outbreak investigation will determine which procedure(s) - environmental improvement, rodent-proofing, and domestic rodent extermination by rodenticides - are to be used and in which order. In an outbreak of plague, the first step in control operations is to use insecticides to kill rat flea before using rodenticides to kill the rats.

Zoonoses

Different routes of transmission to human are possible, as follows:

- direct;
- thought arthropods and rodents;
- thought food and the environment.

Direct transmission is mainly an occupational risk of veterinary personnel, farmers, and hunters, and may be more frequent in areas of poor hygiene. Control measures for outbreak resulting from direct contact with animals vary, depending on the diseases and circumstances (15).

Food-borne diseases

Food-borne diseases may be divided into intoxication (food poisoning) and infections. Outbreaks are most frequently caused by noroviruses, rotaviruses, Salmonella, Clostridium perfringens, Staphylococcus aureus, Bacillus cereus, Campylobacter, Escherichia coli, Clostridium botulinum, and Yersinia enterocolitica. However, in many cases, the origin of food-borne outbreaks remains unknown. Viruses may be more frequently involved than indicated by present data.

Measures include: elimination of contaminated food and prevention of extension of the outbreak by withdrawal of suspected food from market, treatment of contacts, identification of infected food handlers, recommendations for good food handling and preparing practice (15-19).

Case study

Outbreak of Q fever among a group of high school students in Slovenia, March-April 2007

Introduction

Infection with Coxiella burnetii causes Q fever which occurs sporadically or in outbreaks in endemic areas (46). Q fever is an underdiagnosed and underreported communicable disease as majority of cases remains asymptomatic or manifest as a nonspecific flu-like febrile illness (47-52). According to ECDC report, there were 958 notified cases of Q fever in year 2005 in European Union, mainly from Germany, France and Spain (53). The number of notified Q fever cases in Slovenia is very low, with incidence rate varying between 0.15-0.25 cases per 100.000 (54).

A group of 33 veterinary students and two teachers contracted a laboratory-confirmed Q fever infection during a training course on a sheep farm in Slovenia in March 2007. The phases of outbreak investigation are described (55).

Outbreak

Outbreak investigation

On 17 April 2007 the Communicable Disease Centre at the National Public Health Institute was informed about a case of Q fever in an 18 year-old student of a veterinary high school. The patient had developed high fever and a severe headache on 30 March 2007. Chest X-rays showed pneumonia. The student reported that her classmates in the same school year had been complaining about similar symptoms.

We suspected that the patient and her schoolmates might have been exposed to a common source of Q fever when attending a training course on a sheep farm (Farm A) located in the south-western part of Slovenia in March 2007. An outbreak investigation was launched involving all third grade students of the veterinary high school.

Epidemiological and clinical data

The patient's high school year has 66 students. As part of their training, 45 of the students spent several hours between 5 and 23 March 2007 on Farm A, together with three teachers. They were trimming sheep's feet, disinfecting wounds, and applying intramuscular vitamins and anti-helmintic injections as a preventive measure to healthy animals in a stable with approximately 500 sheep. Parturition time of the sheep on Farm A was from end of January to beginning of March this year.

We interviewed all individuals who had been at Farm A, and tested them for Q fever. In addition, 20 students who had not participated in the training course in March were included as a control group. One student from the control group had spent a short time on Farm A in autumn 2006. All interviewed individuals (68 altogether) had been in contact with different domestic animals at home and during school training courses at several locations.

Among 48 exposed individuals, there were 34 (71%) with high fever (38°C or more) with or without a headache. Four individuals (8%) had serious headaches only, but were not sure about fever, and three individuals (6%) reported symptoms of a common cold. Seven individuals (15%) were asymptomatic. The first person to develop symptoms of high fever and headache started to feel ill on 20 March 2007.

Among the 20 control subjects (including the student who had spent a short time on the sheep farm in autumn 2006), three students reported having a prolonged cough, and one had symptoms of a common cold. Sixteen (80%) students were completely asymptomatic.

Laboratory investigation

Serum samples were collected from 63 individuals (93% of the 68 interviewees) and sent to the Institute of Microbiology and Immunology, Medical Faculty in Ljubljana. They were tested by indirect immune-fluorescence (FOCUS diagnostics) for the presence of IgG and IgM antibodies to C. burnetii phase I and II antigens.

A laboratory-confirmed case of acute Q fever was defined as an individual with IgM titres higher than 1:16 and/or IgG titres higher than 1:256.

Overall, the results confirmed acute Q fever in 36 individuals (57% of the 63 tested). Serum samples from 26 individuals (41%) were negative and the test result of one sample was inconclusive.

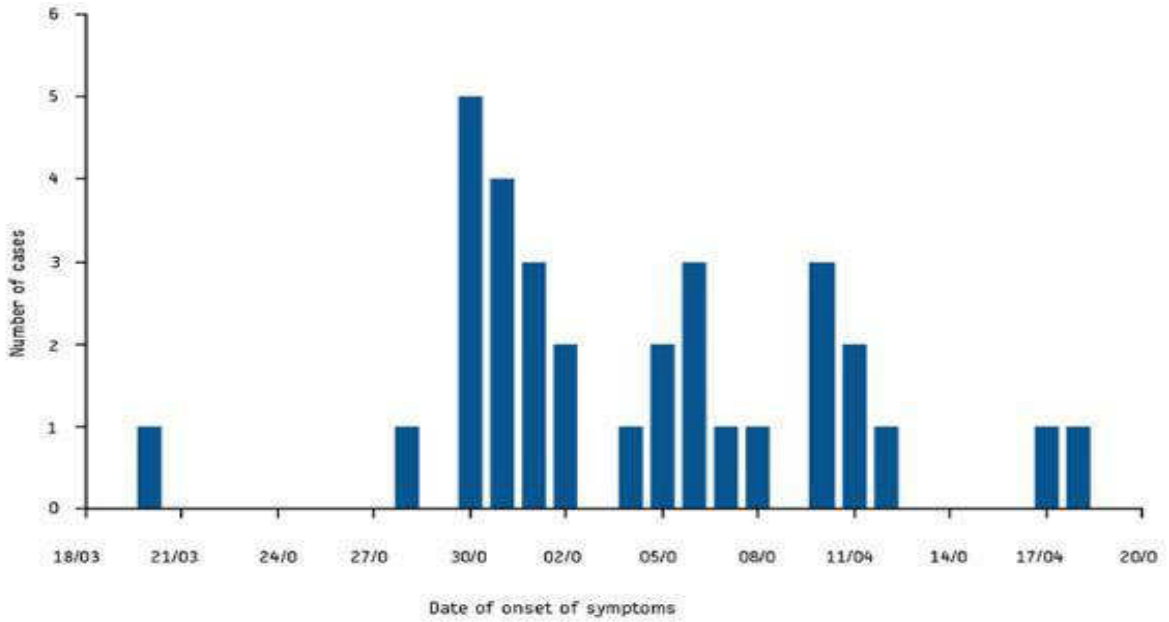
Of a total of 48 individuals who were exposed on Farm A in March 2007, 44 were tested and 35 of those (80%) were seropositive. They are shown in the Figure by date of onset of symptoms. Eight exposed individuals (18% of 44 tested) were not infected, and one had an inconclusive result. In the control group, none was seropositive, except for one student who visited the sheep farm in autumn 2006. The serology results of exposed versus non-exposed individuals are summarised in the Table 1.

Table 1. Laboratory confirmed cases of Q fever in individuals exposed or not exposed on a sheep farm

	Exposed	Not exposed	Total
Seropositive	36	0	36
Seronegative	8	18	26
Inconclusive	1	0	1
Not tested	4	1	5
	49	19	68

FIGURE

Laboratory-confirmed Q fever cases by onset of symptoms, Slovenia, March and April 2007 (n=32)*.



* All cases shown had visited Farm A in March 2007. The date of onset of symptoms was known for only 32 of the 35 cases.

Among the 36 laboratory-confirmed cases (35 exposed in March 2007 and one exposed in autumn 2006), four were asymptomatic, while the other 32 (89%) suffered from high fever, headache, chills, muscle aches, sweating and nausea. Twenty-five patients (69%) consulted a physician, and three (8%) developed radiologically confirmed pneumonia. Although 13 cases (36%) were given antibiotic treatment, only six (16%) received adequate antibiotic therapy with doxycycline, quinolones or macrolide antibiotic. At the time the treatment was initiated, the laboratory results indicating Q fever infection were not known.

The data analysis showed a significant statistical correlation between a positive serological test for C. burnetii and attendance of the training course on the sheep farm in March 2007 (p value <0,001). There was no significant statistical correlation between a positive test and attendance of school training courses in other places or contact with domestic animals at the students' homes. We concluded that the sheep farm was the source of the Q fever outbreak.

Outbreak control measures

The first Q fever outbreak among humans in Slovenia was mentioned in 1954. Two outbreaks were described in 1985. The source of one was contact with a sheep herd; the other occurred among workers at a tannery, who were probably exposed through contact with sheep hides. Further outbreaks due to contact with infected sheep were reported in 1991 and 1992. Between 1996 and 2005, between zero and five Q fever cases were notified in Slovenia.

The Communicable Disease Centre at the National Public Health Institute was collaborating with regional epidemiologists from Ljubljana and Koper, and with the veterinary high school and faculty of Ljubljana on the outbreak investigation. Further collaboration with the Department for Infectious Diseases and the Health Inspectorate at the Ministry of Health, and the Veterinary Office at the Ministry of Agriculture resulted in a cascade of public health and veterinary measures.

Public health measures

Public health measures were as follows:

- suspension of all student training courses and visits to the sheep farm;
- ban on selling dairy products from the sheep farm;

- improvement of sanitary conditions on the sheep farm;
- serological testing of farm employees;
- workplace risk reassessment (farm employees, forestry workers in the vicinity);
- serological testing of students who had been trained on the sheep farm in 2007;
- clinical and serological follow up of seropositive human cases.

Veterinary measures

- serological testing of animals at the sheep farm (as found later, 60% seropositive animals);
- re-introduction of Q fever monitoring in small and large ruminants proposed by the Veterinary Chamber;
- vaccination of animals considered.

The regions endemic for Q fever have never been determined in Slovenia. A joint research project is planned by public health and veterinary scientists to investigate the burden of disease in animals and humans, to determine endemic areas of Q fever in order to develop a sufficient basis for workplace risk assessments, and to determine the risk for the general population.

Exercises

Task 1

Carefully read the part on theoretical background of this module. Critically discuss outbreak investigation with your colleagues.

Task 2

From domestic (e.g. Biomedicina Slovenica, and COBISS-Cooperative Online Bibliographic System of Slovenia in Slovenia), and/or international bibliographic data-bases (e.g. Medline, PubMed) find out the reports of outbreak investigation from your country.

Task 3

After finding an example, discuss the crucial steps of outbreak investigation. Analyze how these steps have been taken in Q fever outbreak presented here.

Task 4

Discuss the strength, limitations, and obstacles in different outbreak investigations.

References

1. Friis R, Sellers TA, Epidemiology for public health practice. An Aspen Publication Inc., Gaithersburg, Maryland, 1999.
2. Baker DJO. Practical Epidemiology. Churchill Livingstone Lmt., London 1973. Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press, 2002.
3. Morton RF, Hebel RJ. A study guide to epidemiology and biostatistics. An Aspen Publication Inc., Rockville, Maryland, 1984.
4. Last JM. A dictionary of epidemiology. Oxford: Oxford University Press, 2001.
5. Premik M. Introduction to epidemiology (in Slovene). Ljubljana: Faculty of Medicine, 1998.
6. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization, 1998.
7. Liu JW, Lu SN, Chen SS, Yang KD, Lin MC, Wu CC, Bloland PB, Park SY, Wong W, Tsao KC, Lin TY, Chen CL. Epidemiologic study and containment of a nosocomial outbreak of severe acute respiratory syndrome in a medical center in Kaohsiung, Taiwan. Infect Control Hosp Epidemiol 2006;27:466-72.
8. Paquet C, Coulombier D, Kaiser R, Ciotti M. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. Euro Surveill 2006;11:212-4.
9. Wang ShX, Li YM, Sun BC, Zhang SW, Zhao WH, Wei MT, Chen KX, Zhao XL, Zhang ZL, Krahn M, Cheung AC, Wang PP. The SARS outbreak in a general hospital in Tianjin, China -- the case of super-spreader. Epidemiol Infect 2006;134:786-91.
10. Porten K, Faensen D, Krause G. SARS outbreak in Germany 2003: workload of local health departments and their compliance in quarantine measures--implications for outbreak modeling and surge capacity? J Public Health Manag Pract 2006;12:242-7.
11. Fox JD. Respiratory virus surveillance and outbreak investigation. J Clin Virol 2007;40 Suppl 1:S24-30.

12. Tan J, Mu L, Huang J, Yu S, Chen B, Yin J. An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. J Epidemiol Community Health 2005;59:186-92.
13. Ho AS, Sung JJ, Chan-Yeung M. An outbreak of severe acute respiratory syndrome among hospital workers in a community hospital in Hong Kong. Ann Intern Med 2003;7;139:564-7. http://www.annals.org/cgi/content/abstract/139/7/564.
14. Boivin G, De Serres G, Hamelin ME, Côte S, Argouin M, Tremblay G, Maranda-Aubut R, Sauvageau C, Ouakki M, Boulianne N, Couture C. An outbreak of severe respiratory tract infection due to human metapneumovirus in a long-term care facility. Clin Infect Dis 2007;44:1152-8.
15. Martinez A, Dominguez A, Torner N, Ruiz L, Camps N, Barrabeig I, Arias C, Alvarez J, Godoy P, Balaña PJ, Pumares A, Bartolome R, Ferrer D, Perez U, Pinto R, Buesa J; Catalan Viral Gastroenteritis Study Group. Epidemiology of foodborne norovirus outbreaks in Catalonia, Spain. BMC Infect Dis 2008;8:47.
16. Outbreaks of Salmonella infections associated with eating Roma tomatoes--United States and Canada, 2004. Can Commun Dis Rep 2005;31:225-8.
17. Smith KE, Stenzel SA, Bender JB, Wagstrom E, Soderlund D, Leano FT, Taylor CM, Belle-Isle PA, Danila R. Outbreaks of enteric infections caused by multiple pathogens associated with calves at a farm day camp. Pediatr Infect Dis J 2004;23:1098-104
18. Arshad MM, Asmar HA, Rahbar MH, Boulton ML, Wells E, Wilkins MJ, Saeed AM. Risk factors for Salmonella oranienburg outbreak in a nursing home in Michigan.J Am Geriatr Soc 2006;54:715-7.
19. Talbot EA, Gagnon ER, Greenblatt J. Common ground for the control of multidrug-resistant Salmonella in ground beef. Clin Infect Dis 2006;42:1455-62.
20. Centers for Disease Control and Prevention (CDC). Dengue outbreak--Federated States of Micronesia, 2012-2013.MMWR 2013;62:570-3.
21. Centers for Disease Control and Prevention (CDC).West Nile virus and other arboviral diseases--United States, 2012. MMWR 2013;62:513-7.
22. Chopra S, Sheevani, Kaur J, Mahajan G. An outbreak of dengue fever in northern India. J Indian Med Assoc 2012;110:751.
23. Hudopisk N, Korva M, Janet E, Simetinger M, Grgic-Vitek M, Gubenšek J, Natek V, Kraigher A, Strle F, Avšic-Županc T. Tick-borne encephalitis associated with consumption of raw goat milk, Slovenia, 2012. Emerg Infect Dis 2013;19:806-8.
24. Chaparro P, Padilla J, Vallejo AF, Herrera S. Characterization of a malaria outbreak in Colombia in 2010. Malar J 2013;12:330.
25. Özbilgina A, Topluoglu S, Es S, Islek E, Mollahaliloglu S, Erkoc Y.Malaria in Turkey: successful control and strategies for achieving elimination. Acta Trop 2011;120:15-23.
26. Yin JH, Yang MN, Zhou SS, Wang Y, Feng J, Xia ZG. Changing Malaria Transmission and Implications in China towards National Malaria Elimination Programme between 2010 and 2012. PLoS One 2013;8:e74228.
27. Spala G, Panagiotopoulos T, Mavroidi N, Dedoukou X, Baka A, Tsonou P, Triantafyllou P, Mentis A, Kyriazopoulou V, Melidou A, Tsiodras S. A pseudo-outbreak of human A/H5N1 infections in Greece and its public health implications. Euro Surveill 2006;11:263-7.
28. Areechokchai D, Jiraphongsa C, Laosiritaworn Y, Hanshaoworakul W, O'Reilly M; Centers for Disease Control and Prevention (CDC). Investigation of avian influenza (H5N1) outbreak in humans--Thailand, 2004. MMWR 2006;55: Suppl 1:3-6.
29. Jones BA, Grace D, Kock R, Alonso S, Rushton J, Said MY, McKeever D, Mutua F, Young J, McDermott J, Pfeiffer DU. Zoonosis emergence linked to agricultural intensification and environmental change. Proc Natl Acad Sci 2013;110:8399-404.
30. Lim LS, Varkey P, Giesen P, Edmonson L. Cryptosporidiosis outbreak in a recreational swimming pool in Minnesota. J Environ Health 2004;67:16-20.
31. Nygard K, Schimmer B, Sobstad O, Walde A, Tveit I, Langeland N, Hausken T, Aavitsland P. A large community outbreak of waterborne giardiasis-delayed detection in a non-endemic urban area. BMC Public Health 2006;6:141. http://www.biomedcentral.com/1471-2458/6/141
32. Beaudeau P, de Valk H, Vaillant V, Mouly D, Ledrans M, Mannschott C, Tillier C. Lessons learned from ten investigations of waterborne gastroenteritis outbreaks, France, 1998-2006. J Water Health 2008;6:491-503.

33. Fong TT, Mansfield LS, Wilson DL, Schwab DJ, Molloy SL, Rose JB. Massive microbiological groundwater contamination associated with a waterborne outbreak in Lake Erie, South Bass Island, Ohio. Environ Health Perspect 2007;115:856-64.
34. Morrow CB, Novick LF. A case exercise in public health preparedness: a community outbreak of influenza-like illness. J Public Health Manag Pract 2005;11:306-10.
35. Balicer RD, Huerta M, Levy Y, Davidovitch N, Grotto I. Influenza outbreak control in confined settings. Emerg Infect Dis 2005;11:579-83.
36. Young LC, Dwyer DE, Harris M, Guse Z, Noel V, Levy MH; Prison Influenza Outbreak Investigation Team. Summer outbreak of respiratory disease in an Australian prison due to an influenza A/Fujian/411/2002(H3N2)-like virus. Epidemiol Infect 2005;133:107-12.
37. Horcajada JP, Pumarola T, Martínez JA, Tapias G, Bayas JM, de la Prada M, García F, Codina C, Gatell JM, Jimenez de Anta MT. A nosocomial outbreak of influenza during a period without influenza epidemic activity. Eur Respir J 2003 Feb;21:303-7.
38. Black M, McAnulty J. The investigation of an outbreak of cryptosporidiosis in New South Wales in 2005. N S W Public Health Bull 2006;17:76-9.
39. Centers for Disease Control and Prevention (CDC). Outbreak of multidrug-resistant Salmonella enterica serotype Newport infections associated with consumption of unpasteurized Mexican-style aged cheese--Illinois, March 2006-April 2007. MMWR Morb Mortal Wkly Rep 2008;57:432-5. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5716a4.htm
40. Lim LS, Varkey P, Giesen P, Edmonson L. Cryptosporidiosis outbreak in a recreational swimming pool in Minnesota. J Environ Health 2004;67:16-20.
41. Dalben M, Varkulja G, Basso M, Krebs VL, Gibelli MA, van der Heijden I, Rossi F, Duboc G, Levin AS, Costa SF. Investigation of an outbreak of Enterobacter cloacae in a neonatal unit and review of the literature. J Hosp Infect 2008;70:7-14.
42. Murray RJ, Pearson JC, Coombs GW, Flexman JP, Golledge CL, Speers DJ, Dyer JR, McLellan DG, Reilly M, Bell JM, Bowen SF, Christiansen KJ. Outbreak of Invasive Methicillin-Resistant Staphylococcus aureus Infection Associated With Acupuncture and Joint Injection. Infect Control Hosp Epidemiol. 2008 Aug 6.
43. Nygard K, Schimmer B, Sobstad O, Walde A, Tveit I, Langeland N, Hausken T, Aavitsland P. A large community outbreak of waterborne giardiasis-delayed detection in a non-endemic urban area. BMC Public Health 2006;6:141. http://www.biomedcentral.com/1471-2458/6/141
44. Meerburg BG, Singleton GR, Kijlstra A. Rodent-borne diseases and their risks for public health. Crit Rev Microbiol 2009;35:221-70.
45. Heyman P, Vaheri A, Lundkvist A, Avsic-Zupanc T. Hantavirus infections in Europe: from virus carriers to a major public-health problem.Expert Rev Anti Infect Ther 2009;7:205-17.
46. Cutler SJ, Bouzid M, Cutler RR. Q fever. J Infect 2007;54:313-8.
47. Van Woerden HC, Mason BW, Nehaul LK, Smith R, Salmon RL, Healy B, et al. Q fever outbreak in industrial setting. Emerg Infect Dis 2004;10:1282-9.

    http://www.cdc.gov/Ncidod/eid/vol10no7/03-0536.htm
48. Sukrija Z, Hamzic S, Cengic D, Beslagic E, Fejzic N, Cobanov D, Maglajlic J, Puvacic S, Puvacic Z Ann N Y Acad Sci 2006;1078:124-8. Human Coxiella burnetii infections in regions of Bosnia and Herzegovina, 2002.
49. Medic A, Dzelalija B, Punda Polic V, Gjenero Margan I, Turkovic B, Gilic V. Q fever epidemic among employees in a factory in the suburb of Zadar, Croatia. Croat Med J 2005;46):315-9.
50. Carrieri MP, Tissot-Dupont H, Rey D, Brousse P, Renard H, Obadia Y, Raoult D. Investigation of a slaughterhouse-related outbreak of Q fever in the French Alps. Eur J Clin Microbiol Infect Dis 2002;21:17-21.
51. Manfredi Selvaggi T, Rezza G, Scagnelli M, Rigoli R, Rassu M, De Lalla F, Pellizzer GP, Tramarin A, Bettini C, Zampieri L, Belloni M, Pozza ED, Marangon S, Marchioretto N, Togni G, Giacobbo M, Todescato A, Binkin N. Investigation of a Q-fever outbreak in northern Italy. Eur J Epidemiol 1996;12:403-8.
52. Parker NR, Barralet JH, Bell AM. Q fever. Lancet. 2006 Feb 25;367:679-88.
53. Annual Epidemiological Report on Communicable Diseases in Europe. European Centre for Disease Prevention and Control. Stockholm, Sweden, June 2007.
54. Surveillance of communicable diseases (in Slovene: Epidemiološko spremljanje nalezljivih bolezni v Sloveniji v letu 2011, CDC report, NIPH, Slovenia: http://www.ivz.si/

55. Grilc E, Socan M, Koren N, Ucakar V, Avsic T, Pogacnik M, Kraigher A. Outbreak of Q fever among a group of high school students in Slovenia, March-April 2007. Euro Surveill 2007;12(29):pii=3237. Available online: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=3237.

Recommended readings

1. Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press, 2002.
2. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization, 1998
3. Ofner-Agostini M, Wallington T, Henry B, Low D, McDonald LC, Berger L, Mederski B; SARS Investigative Team, Wong T. Investigation of the second wave (phase 2) of severe acute respiratory syndrome (SARS) in Toronto, Canada. What happened? Can Commun Dis Rep 2008;34:1-11.
4. Areechokchai D, Jiraphongsa C, Laosiritaworn Y, Hanshaoworakul W, O'Reilly M; Centers for Disease Control and Prevention (CDC).Investigation of avian influenza (H5N1) outbreak in humans--Thailand, 2004. MMWR Morb Mortal Wkly Rep 2006;28;55 Suppl 1:3-6.
5. Gilsdorf A, Kroh C, Grimm S, Jensen E, Wagner-Wiening C, Alpers K. Large Q fever outbreak due to sheep farming near residential areas, Germany, 2005. Epidemiol Infect 2008;136:1084-7.
6. Porten K, Rissland J, Tigges A, Broll S, Hopp W, Lunemann M, van Treeck U, Kimmig P, Brockmann SO, Wagner-Wiening C, Hellenbrand W, Buchholz U. A super-spreading ewe infects hundreds with Q fever at a farmers' market in Germany. BMC Infect Dis 2006;6:147.
7. Brouqui P, Badiaga S, Raoult D. Q fever outbreak in homeless shelter. Emerg Infect Dis 2004;10:1297-9.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Evaluation in Health Promotion**
Module: 2.24	ECTS: 0.25
Author(s), degrees, institution(s)	Alexandra Cucu, MD, PhD, Senior lecturer University of Medicine and Pharmacy "Carol Davila" Bucharest, Director National Centre for Health Promotion and Evaluation – National Institute of Public Health, Bucharest, Romania
Address for correspondence	Alexandra Cucu, MD, PhD National Institute of Public Health Bucharest Dr. Leonte str.; no. 1-3; sector 5; 050463 Bucharest, Romania Phone : +40 21 318 36 20/ ext. 181 Mobile: +40 723 93 05 79 Email: alexandracucu2003@gmail.com
Key words	Health promotion, health promotion interventions, programme evaluation.
Learning objectives	After completing this module students and public health professionals will be able to: • understand the conceptual framework of health promotion and its evaluation; • increase knowledge on basic principles of evaluation; • improve knowledge on health promotion evaluation; • becoming aware of the necessity of health promotion intervention evaluation; • be able to evaluate a health promotion intervention.
Abstract	Evaluation of health promotion intervention is a systematic examination and assessment of input, process and outcomes of a health promotion intervention in order to produce information for further improvement. Due to complexity of health promotion interventions, several methodological and practical issues have to be clarified from the beginning of the processes. In this context, based on the literature review, and the model suggested by Springett and al., an eight step evaluation framework and its principles is described as a general planning guide for the evaluation process. Some important methodological issues and challenges specific for each step are further detailed. Also, critical points and difficulties are briefly presented. The whole approach contributes to enhance understanding of methodology and importance of evaluation as part of health promotion interventions cycle.
Teaching methods	Teaching methods include: • introductory lectures related to health promotion concept and its understanding; • distribution and discussion of relevant literature on health promotion and its evaluation and best cases examples; • guided discussion on general health promotion interventions and their evaluation; • small group evaluation for a health intervention from the best cases examples; • distribution of topics for seminar papers.
Specific recommendations for teachers	Specific recommendations: • ¾ lectures; ¼ discussions; • facilities equipment available; • training materials elaborated and distributed; • best cases health promotion interventions presentations.
Assessment of Students	Assessment of students: • multiple-choice questionnaire for theoretical aspect • presentation of evaluation papers

EVALUATION IN HEALTH PROMOTION

Alexandra Cucu

Introduction

Generally, there are more than 100 specific types of evaluation, each of them appropriate for specific purposes, from the very broad perspective of Green (1) as "comparison with an object of interest against a standard of acceptability" to the programme/project more specific evaluation.

Health promotion evaluation shares many issues common to evaluation in general, but due to specificity of community health interventions, raise many methodological difficulties. In order to be able to set the most appropriate approach for health promotion intervention a clear definition of concepts and understanding of health promotion and evaluation should be done from the beginning.

Health promotion concept

Even though rather new, the health promotion concept and approach due to the recent social, economic, demographic and technological developments has evolved, continuously broadening its senses. Any overview of health promotion definitions should start from Marc Lalonde (2) approach, who, since 1974, on his document "A new perspective on health of Canadians" identifies health promotion as a key *strategy "aiming at informing, influencing and assisting both individuals and organizations so that they will* accept more responsibly and be more active in matters affecting mental and physical health" emphasizing both on information and assistance rolls at individual and organizational level.

Few years later, the U.S. Department of Health Education and Welfare definition (3), contributes to widening the modern understanding of the HP as "a combination of heath education and related organizational, political and economic programs designed to support changes in behavior and in the environment that will improve health" (4), stressing the integrated, multilevel approach of the health promotion intervention and also defining the ultimate goal of that process, the health improvement.

The recent conceptualization of the HP, corresponding to the WHO updated definition contained in the Ottawa Charter for Health Promotion (1986), (5) "the process of enabling people to increase control *over the determinants of their health and thus to improve their health"* is focusing in another core principle of health promotion, empowering, as a tool for individual and community action for health improvement.

Finally, the new European Action Plan for Strengthening Public Health Capacities and Services, adopted by the WHO Euro member states reiterate health promotion as one of the 10 core, essential public health operations of public health. The new operational perspective proposed for the health promotion activities avenue in the next decade is restraint to preventive actions focusing on behavioral risk factors.(6)

Even though different, stressing more on finalities, partnership or process features, all those definitions are build on common values as equity, participation and empowering, oriented to the same, consistent goal over decades, improving health and wellbeing at individual and social level.

A framework for evaluation of health promotion interventions

Theoretician's definition on evaluation as "systematic examination and assessment of features of programme or other intervention in order to produce knowledge that different stakeholders can use for a variety of purposes" (7) gives us only an introduction on the complexity of this process. According to the specialist opinion "many methodological issues are associated with evaluation in health promotion, above and beyond the difficulties of program evaluation". That's why an approach based on the logic model framework for programme evaluation, figured below, could act as the simplified necessary background for describing and understanding the specificities and difficulties of health promotion evaluation (8). Its general diagram includes the following components:

Figure 1. Logic model programme evaluation diagram

The significance of terms used is the following:

Resources/Inputs - resources/inputs include the human, financial, organizational, and community resources available for doing the work;

• **Activities -** activities are what the intervention does with the resources in order to reach the intended results;

• **Outputs** - the outputs are the direct products of program activities and may include the concrete results of activities designed and carried out in order to generate the desired impact of the programme (guidelines elaboration, training session for health promoters, information campaigns);
• **Outcomes** - outcomes are the specific changes in program participants' behavior, knowledge, skills, status and level of functioning. Usually attainment of such changes is a long term process. Consequently, short-term outcomes might be attainable within 1-3 years, while longer-term outcomes should be achievable within a 4-6 year timeframe;
• **Impact** - Impact is the fundamental intended change occurring in individuals, organizations, communities or systems as a result of program activities. It often occurs after the programme/intervention ends.

In applying this simplified approach, the health promotion specific attributes, mainly complexity, dependence of political and social values and context, diversity of results and difficulty to measure outcomes as participation and empowering, should be taken in consideration and carefully addressed. That's why, in order to set the frame and instruments for health promotion evaluation, several methodological and practical issues have to be clarified when evaluating a programme.

The main methodological aspect is related to difficulty to determine the relationship between the intervention and their associated, sometimes synergic, results, in the context of long term effects and multilevel character of health promotion interventions. In addition, reaching best evidences of impact of the interventions, often requiring use of both objective and subjective measurements and ensuring the appropriate level of precision is another issue to be solved.

Other practical aspects are related to difficulties to involve, in this compulsory participatory process, of all relevant stakeholders, objective evaluators and also the beneficiary of the intervention. Another key issue is clear definition of the object of evaluation, one single intervention or a package of activities and their respective associated outcomes. The challenge to deal with results in terms of cost-effectiveness or cost benefit is another difficulty in evaluating the results.

Consequently, if the previously mentioned health promotion activities features are addressed, the general health promotion evaluation framework elaborated, in accordance with specialist opinion, (9) should be based and respond to the following principles:

• be applicable in all evaluation process, but ensure that the most appropriate method is use for the programme, intervention or policy being assessed;
• be consistent with health promotion principles, in particular empowering individuals and communities, emphasizing participation, focus on collective as well as individual accountability;
• be flexible in its application, able to respond to changing circumstances;
• cover all stages of the evaluation process, from setting the agenda to using the results;
• apply to at all levels of evaluation from punctual intervention, projector or programme

Bearing in mind all the above mentioned features, the generic programme evaluation process is, in principle, consisting of several steps, as following:

- Step 1 - **Describing the evaluated programme.** The first step is a crucial one, requiring participation, involvement and commitment, influencing not only the evaluation process but the implementation of its results. This stage includes the team selection for conducting evaluation and sometimes requires a task force team for support of process progression to be established. It also supposes the clarification of the aim and purpose of the evaluation. This steps also suppose collecting baseline information on the evaluated intervention/programme in terms of mandate, aims, objectives, target population, activities and expected results, implementation planned schedule, context and links with other initiatives.
- Step 2 - **Identification of issues of concern.** This step is a major one for clarification of the substance of the evaluation. It consists in formulation of the evaluation question that will establish the aspects addressed in accordance with evaluation purpose and stakeholders point of view. Those should address both the resources, activities and its effects. Also aspects related to intervention implementation process and the related measurable impacts should be clearly identified. This stage represents the basis for the clarification of information needs and indicators to be further elaborated.
- Step 3 - **Designing of the data collection process.** Step tree is related to designing the data collection process. It is a decisive, mostly methodological, step in progression of the evaluation process. It includes decisions on type of evaluation, methods and indicators to be produced in order to respond to the previously formulated questions. It requires involvement and participation of the stakeholders for selecting the most relevant information providing answers for the questions previously selected, ensuring effective evaluation. It has to establish the paradigm and criteria for goals achievement specific to the evaluated intervention. Also, decisions should be taken on what data should be collected. This is closely linked to the agreed methodological approach, quantitative data being usually preferably to qualitative one. But often, due to complexity of health promotion

outcomes qualitative data, meaningful soft tools have to be produced and used. Another important issue is to establish the necessary appropriate level of aggregation of the information, being known that data aggregated for individuals are not always appropriate for assessing the community impact of an intervention at community level (10). Selected measures should reflect both process and outcome, and for the last one effect should be explored both at individual and community/systemic level.

- Step 4 - **Data collection.** The fourth stage is data collecting according to the format and requirements agreed before. This is a process also complicated due to issues as confidentiality, ethical approach and selection of the target group beneficiary of the intervention. Participation and involvement of stakeholders and beneficiaries representatives is necessary for obtain reliable and accurate data for the studied intervention.
- Step 5 - **Data analysis and interpretation.** Step five is dedicated to data analysis and interpretations are another critical step. As often qualitative data are used and their quality is often beyond influence, analysis of data should be carefully made in order to keep and transmit the correct message in an understandable and significant format for the beneficiary of the evaluation. It gives added value to the evaluation process trough translating the technical results in an easy accessible format, adequate for the purpose of evaluation itself.
- Step 6 - **Formulation of recommendations.** The sixth step, succeeding to data analysis, is formulation of recommendation. This includes also clarification of implication of the findings and their implementation. It is the primary mode for valorizing the results of the evaluation made. Stakeholder's involvement during this stage will guarantee the adequacy and feasibility of the recommendation and also their future implementation.
- Step 7 – **Dissemination of the results.** On the seventh stage the results of evaluation should be effectively and n*ot "ad hoc" disseminated* (11). It is a step that must be systematically designed and planed, according to a dissemination plan, in order to maximize the use of results of the evaluation process. It should be done, consistent with the previous steps, with participation and involvement of all actors, stakeholders and beneficiary together. Dissemination should clearly transmit information on the scope, team, methods, questions, results and recommendation of the evaluation process. Proper carried out, with the target audience mobilized it represents one of the moments where information can be a powerful tool in empowering communities and individuals (12,13) ensuring in this way the success of the evaluation made. There are opinions that, for this stage, efforts and resources should be devoted as much as for the whole process. In fact dissemination itself is not only the trigger of improvement and implementation of findings and the recommendation for the policy/intervention evaluated but also represents a model for similar initiatives evaluation or even improvement without an explicit evaluation.
- Step 8 - **Intervention.** The last step, the eighth is the intervention one. It consists in implementation of results and recommendation of the evaluation. It requires identification of resources and tools for proposed changes of the evaluated health promotion interventions and should be done in an articulated systematic manner according to a specific action plan. It is the starting point for the next evaluation process of the health promotion intervention and it contribute to integration of evaluation in health promotion practice. These steps, adapted to the features of the intervention to be evaluated, represent the backbone of a generic health promotion evaluation process. If the key challenges, represented by the correct identification of evaluation questions, decision on the design, outcome measures, and adequate data analysis are met, evaluation can be conducted in a systematic, reproductive way.

Instead of a conclusion

A logical answer to address the question on what is the most appropriate methodology to be used for health promotion interventions would be the design of a comprehensive model, adequate for all purposes and interventions. If so, the principles and the logical steps of the evaluation framework presented above are largely applicable. The rest, concerns the role of the evaluation team and their art to involve and motivate stakeholders and beneficiary participation in the process.

To conclude this overview, several features of health promotion evaluation should be always keep in mind when planning such an approach. The most important aspects are as follows:

- health promotion evaluation is a process that requires systematic planning due to complexity of the evaluated theme (12,13);
- it is an ongoing process which adapt and adjust projects and programs to better fit to proposed end point in real life of implementation;
- it requires good evaluators able of logical thinking, ethical approach, excellent communication and interpersonal skills as well as research and conceptualization skills (14,15);

- it is strongly participatory process, stakeholders' involvement during all stages being crucial for progression of evaluation and its added value for the intervention future;
- as a systematic information feedback mechanism, it is necessary for all health promotion intervention, allowing adjustments for reaching the proposed goals (15,17).

Concluding, it is clear that no matter how systematically rigorous is the evaluation process planed and conducted; its results are strongly depending on the team, skills and quality of the evaluators and their capabilities to lead the process in order to reflect the complexity of the intervention and to ensure all stakeholder and beneficiary participation.

In conclusion, well designed and carried out evaluation contributes not only in improving of the evaluated health promotion intervention but also in developing networks and contacts, creating bridges between practitioners, beneficiary and decision makers, increasing the impact, support and participation for other health promotion activities.

Exercises

Task 1

Based on the concepts and steps described above, design the evaluation tools necessary for carrying out an evaluation process for a health promotion intervention.

Task 2

Using the tools elaborated in the previous exercise, design an evaluation program for one of the health promotion school based-interventions you have being exposed in your recent past.

References

1. Green LW, Krauter MM. Health promotion planning: en educational and environmental approach, 2nd ed. Mountain View: Mayfield, 1991.
2. Lalonde M. "A new perspective on health of Canadians: a working document. Ottawa: Canada Information, 1974.
3. Green LW. National Policy in the promotion of health. Internat J Educat, 1979;22:161-8.
4. Green LW and Iverson DC. School Health Education. Annual Rev Pub Health, 1982;3:321-38.
5. WHO, Ottawa Charter for Health Promotion, Health promotion. 1986.
6. WHO, European Action Plan for strengthening public health capacities and services in EuropeWHO Regional Office for Europe, 2012
7. Irving Rotman, et al, editors. Evaluation in health promotion: principles and perspectives. Copenhagen: WHO regional publications, European series No. 92, 2001.
8. Kellogg WK. Foundation Evaluation Handbook. Publisher, 1998.
9. Springett J, et al. Towards a framework for evaluation in health promotion: the importance of the process. J Contemp Health, 1995;2:61-5.
10. Flay BR. Efficacy and effectiveness trials in the development of health promotion programs. Prevent Med, 1986;15:451-74.
11. Potvin L. Methodological challenges in evaluation of dissemination programs. Canad J Pub Health, 1996;87:79-83.
12. Donner A, et al. Randomization by cluster: sample size requirements and analysis. Am J Epidemiol, 1981;114:906-14.
13. Potvin L, Haddad S, Frohlich S. Beyond process and outcome evaluation - a comprehensive approach for evaluating health promotion programmes. In: Evaluation in health promotion: principles and perspectives. Copenhagen: WHO regional publications, European series No.92, 2001. p.45-56.
14. Nutbeam D, et al. Evaluation in health education: a review of progress, possibilities and problems, J Epidemiol Commun Health, 1990;44:83-9.
15. Mathison S. Rethinking the evaluator role: partnership between organizations and evaluators. Evaluation and planning 1994;17:299-304.
16. Everitt A. Values and evidence in evaluating community health. Crit. Pub Health, 1995;6:56-65.
17. Koplan, J. P and all, Framework for Program Evaluation in Public Health USA Centers for Disease Control and Prevention.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Violence against women: measurement and indicators**
Module: 2.25	ECTS (suggested): 0.5
Author(s), degrees, institution(s)	Bosiljka Djikanovic, MD, PhD, Associate Professor Institute of Social Medicine, School of Medicine, University of Belgrade Dr Subotica 15 11000 Belgrade Serbia and Montenegro
Address for correspondence	Bosiljka Djikanovic, MD Institute of Social Medicine, School of Medicine, University of Belgrade Dr Subotica 15 11000 Belgrade Serbia and Montenegro E-mail: bosiljka.djikanovic@mfub.bg.ac.rs Phone: +381 11 2643 830 Fax: +381 11 2659 533
Keywords	Indicators, measurement, violence against women.
Learning objectives	After completing this module students and public health professionals should be able to: • understand the magnitude and importance of violence against women (VAW) as a global public health problem; • identify the most important international documents that address violence against women; • realize why methods for measuring VAW are of the critical importance for obtaining reliable estimates of prevalence and incidence of violence against women; • explain what consists the gold standard for measuring VAW as well as ethical and safety recommendations for VAW research; • specify different aspects of proposed indicators of VAW.
Abstract	Violence against women (VAW) is a world-wide phenomenon, which still exists in almost all cultures and countries. It has been recognized as a global public health problem of the highest priority. Among the most important international documents that addressed VAW and expressed commitment to end violence against women are The Convention on the Elimination of All Forms of Discrimination against Women (CEDAW, 1979); Beijing Platform for Action (1995), United Nations Millennium Declaration (2000), and many other. In order to accurately estimate the incidence and prevalence of VAW, several methodological approaches were used in recent years, such as the WHO Multi-country study on women's health and domestic violence and the International VAW Survey (specialized surveys), or generic DHS surveys with the module on domestic violence. These surveys were generating different disclosure rates, and the methodology developed by the World Health Organization is considered as a gold standard for reliable measurement of this phenomenon, while complying with ethical recommendations that maximize women's safety and well being, at the first place. The United Nations urged member countries to monitor VAW in their populations, in terms of women's age, relationship with the perpetrator, type of violence, severity, frequency, and time period (lifetime or during the last 12 months), forming a set of indicators that would allow mutual comparisons and follow up national progress toward achieving zero tolerance for VAW.
Teaching methods	Teaching methods include lectures, case studies, interactive group discussions, seminars.
Specific recommendations for teachers	This module should be 0.5 ECTS, of which two thirds are lectures and guided discussions under supervision, while one third should be allowed for individual work, i.e. searching available literature, writing seminar paper and preparing for oral presentation and discussion.
Assessment of students	Quality of seminar paper and delivered oral presentation.

VIOLENCE AGAINST WOMEN: MEASUREMENT AND INDICATORS

Bosiljka Djikanovic

"There is one universal truth, applicable to all countries, cultures and communities: violence against women is never acceptable, never excusable, *never tolerable."*

United Nations Secretary-General, Ban Ki-Moon (2008)

Violence against women

Violence against women (VAW) is a world-wide phenomenon, which still exist in almost all cultures and countries. As defined by United Nations, violence against women means any act of gender-based violence that results in, or is likely to result in, physical, sexual or psychological harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or private life (1). The most prevalent form of VAW is violence perpetuated by women's current or former male intimate partners. Intimate partner violence (IPV) happens in more than 75% of all cases of violence against women (2). IPV is defined as any act or omission by a current or former intimate partner, which negatively affects the well-being, physical or psychological integrity, freedom, or right to the full development of a woman (1). IPV might have many facets that differ in terms of type (physical, sexual, psychological, emotional, economic violence), severity (moderate, severe violence), frequency (once, twice, many times) and time period (during the last 12 months or earlier at lifetime). The most recent review of global and regional VAW prevalence studies conducted by World Health Organization used following definitions and they are presented in Table 1 (3).

Violence against women is often referred to as a gender-based violence, as it reflects historical inequalities between women and men, and corresponding discrimination and abuse of women, which is fundamental violation of women's human rights that is absolutely unacceptable. According to the latest estimates, one third of women and girls globally have experienced violence against them, with serious consequences for their physical and mental health, and an overall participation in society (3). It is a global public health problem that in recent years has received a growing recognition among a large number of international and national actors.

Table 1. Working definitions of different forms of exposure to intimate partner violence

TERM	DEFINITION
Intimate partner violence	Self-reported experience of one or more acts of physical and/or sexual violence by **a current or former partner** since the age of 15 years. **Physical violence** is defined as: being slapped or having something thrown at you that could hurt you, being pushed or shoved, being hit with a fist or something else that could hurt, being kicked, dragged or beaten up, being choked or burnt on purpose, and/or being threatened with, or actually, having a gun, knife or other weapon used on you. **Sexual violence** is defined as: being physically forced to have sexual intercourse when you did not want to, having sexual intercourse because you were afraid of what your partner might do, and/or being forced to do something sexual that you found humiliating or degrading.
Severe intimate partner violence	Is defined on the basis of the severity of the acts of physical violence: being beaten up, choked or burnt on purpose, and/or being threatened or having a weapon used against you is considered severe. Any sexual violence is also considered severe.
Current intimate partner violence	Self-reported experience of partner violence in the past year.
Prior intimate partner violence	Self-reported experience of partner violence before the past year.

Source: adapted from WHO, 2013 (3).

This article presents an overview of the most important global documents, conventions, declarations and agreements related to violence against women; methods for measuring women's experience with violence during lifetime and within the last year, as well as a set of indicators that are proposed for monitoring the phenomenon and evaluating the impact of large scale programs and interventions aimed to combat violence against women and establishing a null tolerance to it in every society.

International documents related to violence against women

CEDAW convention

The Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) was adopted in 1979 by the UN General Assembly and it is often described as an international bill of rights for women (4). Consisting of a preamble and 30 articles, it defines what constitutes discrimination against women and sets up an agenda for national action to end such discrimination.

The Convention defines discrimination against women as *"...any distinction, exclusion or restriction made on the basis of sex which has the effect or purpose of impairing or nullifying the recognition, enjoyment or exercise by women, irrespective of their marital status, on a basis of equality of men and women, of human rights and fundamental freedoms in the political, economic, social, cultural, civil or any other field"* (4).

By accepting the Convention, States commit themselves to undertake a series of measures to end discrimination against women in all forms, including:

- to incorporate the principle of equality of men and women in their legal system, abolish all discriminatory laws and adopt appropriate ones prohibiting discrimination against women;
- to establish tribunals and other public institutions to ensure the effective protection of women against discrimination; and
- to ensure elimination of all acts of discrimination against women by persons, organizations or enterprises.

CEDAW convention is an important document that implicitly addresses violence against women, as this phenomenon is a demonstration of historical inequalities between women and men, and women's discrimination. Periodical national reporting every four years (and/or whenever CEDAW committee requests) is an obligation of countries that signed up and ratified The Convention. These reports, which may indicate factors and difficulties in implementation, are forwarded to the CEDAW for its consideration. Reporting to The Committee is very important mechanism for assuring that governmental policies and actions are launched and implemented toward eliminating discrimination of women.

Beijing Platform for Action

The United Nations Fourth World Conference on Women was held in Beijing, China, in September 1995, where the well known Beijing Platform for Action was launched (later in the text: The Beijing Platform) (5). This document proposed strategic objectives of the relevance for violence against women. They are related to taking integrated measures to prevent and eliminate violence against women, and to study causes and consequences of violence against women and the effectiveness of preventive measures (5).

The Platform clearly states that violence against women is an obstacle to the achievement of the objectives of equality, development and peace. Violence against women is a manifestation of the historically unequal power relations between men and women, which have led to domination over and discrimination against women by men and to the prevention of women's full advancement (5). Violence against women is derived essentially from cultural patterns, in particular the harmful effects of certain traditional or customary practices and all acts of extremism linked to race, sex, language or religion that perpetuate the lower status accorded to women in the family, the workplace, the community and society. Women who are particularly vulnerable to violence are women belonging to minority groups, indigenous women, refugee women, women migrants, including women migrant workers, women in poverty living in rural or remote communities, destitute women, women in institutions or in detention, female children, women with disabilities, elderly women, displaced women, repatriated women, women living in poverty and women in situations of armed conflict, foreign occupation, wars of aggression, civil wars, terrorism, including hostage-taking.

This platform recommends that Governments and other actors should promote an active and visible policy of mainstreaming a gender perspective in all policies and programmes so that before decisions are taken an analysis may be made of their effects on women and men, respectively. It proposes a set of operational actions to be taken at national levels, according to the strategic objectives that are implicitly relevant for violence against women (5).

Millennium Declaration and Millennium Development Goals

United Nations Millennium Summit in September 2000 was the largest gathering of world leaders in history when they adopted the UN Millennium Declaration, expressing commitments to a new global partnership to reduce extreme poverty and setting out a series of time-bound (till 2015) development targets that are known as the Millennium Development Goals (MDGs) (6). The MDGs are quantified targets for addressing extreme poverty in its many dimensions-income poverty, hunger, disease, lack of adequate shelter, and exclusion, while promoting gender equality, education, and environmental sustainability. MDGs embedded basic human rights i.e. the rights of each person on the planet to health, education, shelter, and security.

Violence against women is a serious treat for achieving all MDGs, especially MDG 3 that is related to gender equality and women's empowerment, and MDG 5 that is focused on improving maternal health. MDGs and VAW are strongly linked: working towards the MDGs will reduce violence against women, and vice versa, preventing violence against women will contribute to achieving the MDGs (7). These links have to be translated into concrete proposal for actions, as depicted in the World Health Organization publication titled Addressing violence against women and achieving the Millennium Development Goals (7). Discrimination of women is deeply enrooted in social and cultural norms, and greater gender equality and empowerment will help many women to avoid violence, whereas change of men's behavior is of critical importance, as stated in this report: *"(...) the violence will never disappear unless men also change their* attitudes and reject violence against women as acceptable behavior in any context, including in the home *(...)"* (7).

Although the targets and indicators under MDG 3 recognize that education, literacy, wage employment and political participation are important indices of women's empowerment, their achievement – in and of themselves – do not directly address violence against women. This has led the Millennium Project Task Force on Education and Gender Equality to suggest that other country-level targets and indicators be included for this MDG. Related to VAW, one of suggested targets is to reduce the lifetime prevalence of violence against women by 50% by 2015, and governments should ensure continuous monitoring of achieving this goal, by conducting adequate surveys and providing reliable statistics on violence against women, as well as data on perpetrators' prosecution and conviction rates.

Along with international conventions, declarations and resolutions, over the past decade there have been considerable advances in the available body of knowledge related to incidence and prevalence of violence against women and its consequences, thanking to well designed methodologies that enable more accurate measurement of this phenomenon.

Measurement of violence against women

A growing number of research evidences related to the incidence and prevalence of intimate partner violence, as a most common form of violence against women, associated risk factors and health consequences, are now available for many countries worldwide (8). This was possible due to the increased knowledge and understanding of the nature of this phenomenon, and developed methodologies that can accurately collect data on this very sensitive issue. Several considerations are deemed very important for measurement of violence against women, in population and elsewhere. First of all, there are ethical recommendations for women's safety that are a must and every study design aimed to collect data on VAW has to comply with these recommendations, making sure that involvement in the study does not elevate women's risk to experience violence (9). In case of population-based survey, it is important that just one woman in household can be interviewed, in order to keep confidential the content of the questionnaire (9). Interview has to be conducted in a private place, by a well trained female interviewer, while prior to field work, external community resources for support to victims and survivors of violence have to be identified, and their contacts available in case of need (9,10). If these resources do not exist in that particular community, it is important and ethical obligation to establish them as a necessary prerequisite for conducting the survey (9).

In order to identify an optimal strategy for measuring VAW in the population, different methodological approaches were tested and assessed in terms of their effectiveness in a disclosure, since women are hesitating to confirm that they have experienced violence, which lead to a great number of falsely negative responses and underestimation of the phenomenon (3). Therefore, it is very important that the study methods carefully considered this aspect and undertaken safety and technical measures that enable disclosure, while respondents still feel comfortable, as well as interviewers (9).

By now, measuring IPV-VAW appeared in several different formats and forms (8). It was conducted as a module on domestic violence nested in larger study (such as Demographic and Health Survey, DHS) or as stand-alone surveys (such as International Violence Against Women survey, IVAW, and World Health Organization Multi-country study on women's health and domestic violence). Further, there are two versions of the sequence of questions related to violence: starting with questions on non-partner violence, followed by questions on intimate partner, and vice versa. The placement of the module within different parts of the survey was also tested.

A consensus is achieved that exposure to violence should be measured by asking direct questions about behaviorally specific acts of physical and sexual violence that women might have experienced in the context of the intimate partner relationship (Table 1). The most often use instrument for it is Conflict Tactics Scale (CTS) that consists of 6 questions related to different forms of physical violence and 3 questions related to experience with sexual violence (11). CTS scale proves its reliability in several different studies conducted worldwide (12). However, when emotional or psychological abuse is concerned, currently there is a lack of consensus what consist a threshold for this type of abuse and it greatly depends on the cultural norms where women live (3).

Below will be depicted the most important characteristics of above mentioned cross-sectional population studies that were recently conducted in households worldwide and whose findings provided valuable estimates of this phenomenon.

WHO-*Multi Country Study on Women's Health and Domestic Violence* is a stand-alone study specialized to collect data on women's health and their lifetime experiences with both partner and non-partner violence (12). Methodology used in this study was developed by World Health Organization, London School of Hygiene and Tropical Medicine (LSHTM), and Program for Adequate Technologies in Health (PATH), and fully complied to the ethical recommendations for researches on violence against women (9,12). In the period between 2000 and 2003, population surveys were conducted in eight countries with different cultural settings: Bangladesh, Brazil, Japan, Namibia, Peru, Samoa, Thailand and Tanzania. In addition, the study was conducted in Ethiopia and Serbia, which all together made a pool of more than 24,000 women worldwide who responded to the study that used the same methodology, and therefore, their findings are mutually comparable (13-16). Face-to-face interviews were conducted by carefully selected and well trained female interviewers with an ongoing supervision and support during the field work (10,12).

Operational definitions of physical and sexual violence were based on the Conflict Tactics Scale (CTS-2), and these questions were placed in the middle of the questionnaire, after questions related to women's health (11,12). This study collected data on non-partner violence as well; sexual abuse during childhood and forced first sexual experience. This study provided reliable estimates of the prevalence and incidence of VAW, corresponding risk factors, health consequences, suicide attempts among women who experienced violence, etc. (13-16). Initial report ended in 15 recommendations to strengthen national commitment and action on VAW, by promoting primary prevention, education system, strengthening the health sector's response, supporting women living with violence, sensitizing criminal justice, undertaking research and enhancing collaboration (12).

Demographic and Health Surveys (DHS) with a module on domestic violence were the source of national estimates of violence against women in at least fifteen countries worldwide (17). The Demographic and Health Surveys (DHS) are carried out at approximately five-year intervals in a range of mainly low- and middle-income countries. These surveys use largely standardized questionnaires and methodologies and cover a range of topics, including demographics; reproductive, maternal and child health; sexual behavior and nutrition. In-country organizations (usually National Statistical offices) are responsible for implementing the surveys, with technical assistance from Macro International and major funding from the US Agency for International Development (USAID) (17). In the late 1990s, a standardized module of questions on domestic violence was developed; it has since been added to the DHS in 27 countries (18). Findings from nine countries (Cambodia, Colombia, Dominican Republic, Egypt, Haiti, India, Nicaragua, Peru, and Zambia) were extensively compared and presented in the often cited report of Kishor and Johnson (2004) (19).

DHS studies also applied some safety measures, but in general, modules on VAW that are added in these general surveys tend to achieve lower disclosure rates and therefore, to underestimate documented incidence and prevalence rates.

The International Violence Against Women Survey (IVAWS) is an international comparative survey specifically designed to target men's violence against women, especially domestic violence and sexual assaults (20). The IVAWS project largely relies on the network, infrastructure and methodology of the International Crime Victimization Survey (ICVS) along with a specific expertise in development of sensitive survey tools for measurement of VAW that is provided by Statistics Canada, who conducted the first targeted survey in 1993. IVAWS are coordinated internationally by HEUNI, the European Institute for Crime Prevention and Control, with inputs from the UN Office on Drug and Crime (UNODC), UN Interregional Crime and Justice Research Institute (UNICRI), and Statistics Canada (20). Participation in this project is on a self-funded basis, and being conducted by independent investigators in each country. The countries participating in this project were Australia, China (Hong Kong), Costa Rica, the Czech Republic, Denmark, Greece, Italy, Mozambique, Poland, Philippines and Switzerland, and the first comprehensive research report was published in 2008 (21). All nine surveys included questions on violence during pregnancy and four contained sufficient additional information to permit estimates to be made of the prevalence of intimate partner violence among women with children (18,20,21).

The National Institute of Justice and the National Center for Injury Prevention and Control of Centers for Disease Control and Prevention (CDC), who jointly sponsored the National Violence Against Women Survey, conducted CDC National Violence Against Women Survey in United States in 1996 (22). This survey was conducted among both women and men, and data were obtained from 8,000 women and 8,005 men who were 18 years of age or older, residing in households throughout the United States. Respondents were asked about (1) their general fear of violence; (2) emotional abuse they had experienced by marital or cohabiting partners; (3) physical assault they had experienced as children by adult caretakers; (4) physical assault they had experienced as adults by any type of perpetrator; (5) forcible rape or stalking they had

experienced by any type of perpetrator; and (6) threatened violence they had experienced by any type of perpetrator. Findings from this survey are available in a well known and often cited report of Tjaden and Thoennes (2000) and other sources of this survey are available at the CDC site (22,23).

Indicators of violence against women

Recognizing that violence against women is an extremely hurtful phenomenon in the contemporary world, the United Nations General Assembly adopted a set of resolutions calling on governments to undertake all necessary activities and measures to eradicate these practices (24). This highest representative body also requested the UN Secretariat to work on adequate tools for measuring the prevalence and incidence of violence against women, to ensure the systematic collection and analysis of data to monitor all forms of violence against women, by developing universal standards and definitions in that respect (24).

The United Nations Secretary-General's campaign "UNiTE to End Violence against Women" was launched and one of the five key outcomes is the establishment in all countries by 2015 of systems for data collection and analysis (25). The development and use of common indicators on violence against women is critical for a full and comprehensive overview of this phenomenon (8,24-26). A recommended set of questions used to collect data on the interim indicators can prevent countries from diverging in their application of surveys on violence against women and facilitate international comparability of results. (8). Therefore, The UN General Assembly specifically requested that the United Nations Statistical Commission, as the apex body of the international statistical system, to identify and define statistical indicators on violence against women (8,24-26). Consequently, the UN Statistical Commission established the group of countries acting as Friends of the Chair and entrusted it with developing indicators and other methodological standards for measuring VAW prevalence (8,24-27). They come up with a comprehensive set of indicators that distinguish age of women; when violence occurred (during a lifetime or within the last 12 months); type of violence (physical, sexual); severity of violence; the relationship with the perpetrator, and frequency (Table 2). Availability of these indicators and their regular monitoring through national surveys are of the critical importance for raising awareness of this phenomenon and translating it into the concrete actions aimed to reduce women's exposure to violence and establishing null tolerance for it in every society.

Table 2. Set of indicators of violence against women, as recommended by the United Nations (25)

1. Total[1] and age-specific[2] rate of women subjected to **physical violence** in the **last 12 months** by severity of violence, relationship to the perpetrator(s) and frequency
2. Total[1] and age-specific[2] rate of women subjected to **physical violence** during **lifetime** by severity of violence, relationship to the perpetrator(s) and frequency
3. Total[1] and age-specific[2] rate of women subjected to sexual violence in the **last 12 months** by relationship to the perpetrator(s) and frequency
4. Total[1] and age-specific[2] rate of women subjected to **sexual violence** during **lifetime** by relationship to the perpetrator(s) and frequency
5. Total and age-specific rate[3] of women subjected to **sexual or physical violence** by current or former **intimate partner** in the **last 12 months** by frequency
6. Total and age-specific rate[3] of women subjected to **sexual or physical violence** by current or former **intimate partner** during **lifetime** by frequency

[1]Proportion of women aged 15 years and over subjected to physical/sexual violence over the total number of women aged 15 years and over. This rate may be presented as a fraction (e.g. 1 out of 10) or as a percentage (e.g. 10 per cent of total).
[2]Age-specific rate: Proportion of women of a certain age subjected to physical/sexual violence over the total number of women of the same age (5-year age groups starting at 15 years of age). This rate may be presented as a fraction (e.g. 1 out of 10) or as a percentage (e.g. 10 per cent of total).
[3]Total and specific rates for intimate partner violence: Proportion of women aged 15 and over subjected to physical or sexual violence by current or former partner over the total number of women aged 15 and over who have or had an intimate partner.

Exercises

Identification and assessment of national VAW studies in population

Task

Students will work individually or in small groups (up to three) in order to identify any available population VAW study that was conducted in their countries. Students should prepare a written report (seminar paper) on the methodology used in that study, and to critically reflect and discuss its advantages and limitations, based on the content learned during this module. Students should also prepare a Power Point presentation and demonstrate understanding of this topic.

References

1. United Nations. Declaration on the Elimination of Violence against Women. UN General Assembly A/RES/48/104, 20 December 1993. Available at: http://www.un.org/documents/ga/res/48/a48r104.htm (Accessed August 25, 2013).
2. Krug EG, Dahlberg LL, Mercy JA, Zwi AB, Lozano R. World report on violence and health. Geneva, World Health Organization, 2002. Available at: http://whqlibdoc.who.int/publications/2002/9241545615_eng.pdf (Accessed August 25, 2013).
3. Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence. Geneva, World Health Organization, Department for sexual and reproductive health, 2013. Available at: http://www.who.int/reproductivehealth/publications/violence/9789241564625/en/index.html (Accessed August 28, 2013).
4. Convention on the Elimination of All Form of Discrimination against Women (CEDAW). Available at: http://www.un.org/womenwatch/daw/cedaw/cedaw.htm (Accessed August 25, 2013).
5. The Fourth World Conference on Women, Beijing, China, 4-15 September 1995. Being Platform for Action. Available at: http://www.un.org/womenwatch/daw/beijing/platform/violence.htm (Accessed August 25, 2013)
6. United Nations. Millennium development goals (MDG). Available at: http://www.un.org/millenniumgoals/ (Accessed August 25, 2013).
7. World Health Organization. Addressing violence against women and achieving the Millennium Development Goals. Geneva, World Health Organization 2005. Available at: http://www.who.int/gender/documents/MDGs&VAWSept05.pdf
8. United Nations Statistics Division. Methodological Overview of Surveys on Violence against Women. Meeting of the Friends of the Chair of the United Nations statistical Commission on Statistical Indicators on Violence against Women. Aguascalientes, Mexico, 9-11 December 2009. Available at: http://unstats.un.org/unsd/demographic/meetings/vaw/docs/Paper1.pdf (Accessed August 31, 2013)
9. World Health Organization Department of Gender, Women and Health. Putting woman first: ethical and safety recommendations for research on domestic violence against women. Geneva, World Health Organization, 2001. Available at: http://whqlibdoc.who.int/hq/2001/WHO_FCH_GWH_01.1.pdf (Accessed August 31, 2013).
10. Jansen H, Watts C, Ellsberg M, Heise L, Garcia-Moreno C. Interviewer Training in the WHO Multy-Country Study on Women's Health and Domestic Violence. Violence Against Women, 2004;10(7):831-849.
11. Straus MS, Hamby SL, Boney/McCoy S, Sugarman D. The Revised Confilict tactics Scales (CTS2). Development of Preliminary Phychometric Data. Journal of Family Issues 1996;17(3):283-316.
12. World Health Organization Multi-country Study on Women's Health and Domestic Violence against Women. Initial results on prevalence, health outcomes and women's responses. Geneva, World Health Organization, 2005. Available at: http://www.who.int/gender/violence/who_multicountry_study/summary_report/summary_report_English2.pdf (Accessed August 31, 2013).
13. Garcia-Moreno C, Jansen AFMH, Ellsberg M, Heise L, Watts CH, on behalf of the WHO Multi-country Study on Women's Health and Domestic Violence against Women Study Team. Prevalence of intimate partner violence: findings from the WHO multi-country study on women's health and domestic violence. The Lancet 2006;368(9543):1260-69.
14. Ellsberg M, Jansen AFMH, Heise L, Watts CH, Garcia-Moreno C, on behalf of the WHO Multi-country Study on Women's Health and Domestic Violence against Women Study Team. Intimate

partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. The Lancet 2008;371(9619):1165-72.

15. Abramsky T, Watts CH, Garcia-Moreno C, Devries K, Kiss L, Ellsberg M, Jansen AFMH, Heise L. What factors are associated with recent intimate partner violence? Findings from the WHO multi-country study on women's health and domestic violence. BMC Public Health 2011;11:109.
16. Devries K, Watts C, Yoshihama M, Kiss L, Schraiber LB, Deyessa N, Heise L, Durand J, Mbwambo J, Jansen H, Berhane Y, Ellsberg M, Garcia-Moreno C, and WHO Multi-Country Study Team. Violence against women is strongly associated with suicide attempts: Evidence from the WHO multi-country study on women's health and domestic violence against women, Social Science & Medicine 2011;73(1):79-86.
17. MacroDemographic and Health Survey's DHS Toolkit of methodology for the MEASURE DHS Phase III project, implemented from 2008-2013. MEASURE DHS/ICF International. Available at: http://www.measuredhs.com/pubs/pdf/DHSQM/DHS6_Module_Domestic_Violence_28March2013_DHSQM.pdf (Accessed August 25, 2013).
18. Devries KM, Kishor S, Johnson H, Stöckl H, Bacchus LJ, Garcia-Moreno C, Watts C. Intimate partner violence during pregnancy: analysis of prevalence data from 19 countries. Reproductive Health Matters 2010;18(36):158-70.
19. Kishor S, Johnson K. Profiling Domestic Violence – A Multi-Country Study. Calverton, Maryland: ORC Macro, 2004. Available at: http://www.measuredhs.com/pubs/pdf/OD31/OD31.pdf (Accessed August 25, 2013).
20. Nevala S. International Violence Against Women Survey (IVAWS). Expert paper prepared for Expert Group Meeting on Violence against women: a statistical overview, challenges and gaps in data collection and methodology and approaches for overcoming them. UN Division for the Advancement of Women, Geneva, Switzerland, 11-14 April 2005. Available at: http://www.un.org/womenwatch/daw/egm/vaw-stat-2005/docs/expert-papers/Nevala.pdf (Accessed August 25, 2013).
21. Johnson H, Ollus N, Nevala S. Violence Against Women: An International Perspective. Springer & HEUNI, 2008
22. Center for Disease Control and Prevention. Intimate Partner Violence: Data Sources. Available at: http://www.cdc.gov/violenceprevention/intimatepartnerviolence/datasources.html (Accessed August 31, 2013).
23. Tjaden P, Thoennes N. Full Report of the Prevalence, Incidence and Consequences of Violence Against Women. Research report. Findings From the National Violence Against Women Survey. Washington, DC: US Department of Justice, National Institute of Justice, 2000. Available at: https://www.ncjrs.gov/pdffiles1/nij/183781.pdf (Accessed August 31, 2013).
24. United Nations General Assembly, sixty-third session. Agenda item 56 (a). Resolution adopted by the General Assembly A/RES/63/155. Intensification of efforts to eliminate all forms of violence against women. 30 January 2009. Available at: http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/63/155&Lang=E (Accessed on August 31, 2013).
25. United Nations Indicators to measure violence against women. Report to the Expert Group Meeting. Geneva, Switzerland, 8-10 October 2007. Available at: http://www.un.org/womenwatch/daw/egm/IndicatorsVAW/IndicatorsVAW_EGM_report.pdf (Accessed Aug 31, 2013).
26. Walby S. Indicators to measure violence against women. Invited paper. Expert Group Meeting on idnicators to measure violence against women. Geneva, Switzerland, 8-10 October 2007. Available at: http://www.un.org/womenwatch/daw/egm/vaw_indicators_2007/papers/Invited%20Paper%20Walby.pdf (Accessed August 31, 2013).
27. Mrkic S (ed.) The World's Women 2010: Trends and Statistics. Chapter 6: Violence Against Women. New York, United Nations Statistics Division, Department of Economic and Social Affairs, 2010, pg. 127-137. Available at: http://unstats.un.org/unsd/demographic/products/Worldswomen/WW_full%20report_color.pdf (Accessed August 31, 2013).

Recommended readings

1. Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence. Geneva, World Health Organization, Department for sexual and reproductive health, 2013. Available at: http://www.who.int/reproductivehealth/publications/violence/9789241564625/en/index.html (Accessed August 28, 2013).
2. World Health Organization Multi-country Study on Women's Health and Domestic Violence against Women. Initial results on prevalence, health outcomes and women's responses. Geneva, World Health Organization, 2005. Available at: http://www.who.int/gender/violence/who_multicountry_study/summary_report/summary_report_English2.pdf (Accessed August 31, 2013).
3. Mrkic S (ed.) The World's Women 2010: Trends and Statistics. Chapter 6: Violence Against Women. New York, United Nations Statistics Division, Department of Economic and Social Affairs, 2010, pg. 127-137. Available at: http://unstats.un.org/unsd/demographic/products/Worldswomen/WW_full%20report_color.pdf (Accessed August 31, 2013).

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Quality of life: Measurement and cultural adaption**
Module: 2.26	ECTS (suggested): 0.5
Author(s), degrees, institution(s)	Zorica Terzic, Assoc. Prof. PhD, MD, MSc. Bojana Matejic, Assoc. Prof. PhD, MD, MSc. Associate Professor at the School of Medicine, University of Belgrade, Serbia
Address for Correspondence	Zorica Terzic, Institute of Social Medicine, School of Medicine, Belgrade University Dr Subotica 15 11000 Belgrade Serbia Tel: +381 11 2643 830 Fax: + 381 11 2659 533 E-mail: zoricaterzic37@gmail.com
Keywords	Health-related quality of life, measuring quality of life, quality of life, quality of life questionnaire.
Learning objectives	After completing this module students and public health professionals should have: • increased their understanding and knowledge of quality of life and measuring quality of life; • increased knowledge about differences among global, generic and specific instruments of quality of life; • increased knowledge about cultural adaption.
Abstract	It is considered that the first theoretical postulates of measurement of Quality of Life (QoL) as well as measurement scales, have been developed in 1845. The term is used in philosophy literature, sociology, geography, health economics, medicine, psychology, and pharmaceutics industry. Quality of life is separated from the term health-related quality of life (HRQL). The term HRQL has become preferred by many whose task is to assess the impact of disease and treatment on the lives of patients. Measuring QoL is important because it is used for decision making especially about non-clinical aspects of diseases, for improvement of the doctor-patient relationship, in discovering of functional and psychological limitations, in choosing the treatment in initial phase of disease, when the efficiency of an applied therapy is temperate (for example, drugs which just modify a disease). Measuring of quality of life and health-related quality of life could be: one-dimensional and multidimensional. Measuring QoL and HRQoL can be: global and specific (specific in relation to disease and in relation to medical treatment).
Teaching methods	Teaching methods include lectures, students' individual work under the supervision of teachers and interactive methods such as small group discussions. Before the introductory lecture, a small exercise could be organized as brainstorming ("What is quality of life? What are the differences between quality of life and health"), in order to increase students' motivation for learning and the interest in the content of the module. After the introductory lecture, students will work individually on comparison of dimensions among global, generic and specific instruments. Students should discuss in small groups the dimensions of the quality of life that are covered by the general and the specific instruments. They would also have opportunity to search through the Internet under the supervision of teachers in order to explore some of the websites concerning QoL and its related questionnaires.
Specific recommendations for teachers	Teachers should be familiar with the process of SF-36 questionnaire analysis, especially standardization procedure and cultural adaptation.
Assessment of students	Multiple-choice questionnaire.

QUALITY OF LIFE: MEASURMENT AND CULTURAL ADAPTION OF QUESTIONNAIRE

Zorica Terzic, Bojana Matejic

Development of quality of life and health-related quality of life

The first theoretical postulates of measurement of Quality of Life (QoL) as well as measurement scales, have been developed two centuries ago, in 1845 (1). During the next 60 years, theoretical knowledge and research on measurement of health status has been conducted, including Health-Related Quality of Life (HRQoL). The term Quality of Life (QoL) was first mentioned by Pigou in 1920 in his book on economy and welfare (2).

After the Second World War, parallel with economic development and increase of life standard, the importance of happiness, satisfaction and well-being as components of quality of life was recognized and, because of that, in the framework of the social sciences emerged the interest for their exploration (3). Beginning with 1960, quality of life becomes the field of public interest first in North America, where the initial research on QoL was conducted by psychologists, sociologists and social gerontologists, trying to evaluate welfare, satisfaction and happiness. The term QoL in the area of public policy was first presented by the American president Johnson in his speech about importance of the society (3). At the same time, in one medical journal, Elkington published an article titled "Medicine and Quality of Life". Medline introduced "quality of life" as a title in 1975, then Index Medicus adopted it as a concept in 1977 and, after that, quality of life was more and more accepted by various scientific bodies (4).

In the period between 1990 and 1994, the number of clinical researches which measured QoL referring to the period before applying specific therapy is extremely increased (5). Such trend is continuing and orientation toward clinical researches is still indicated.

Historically, the term HRQoL and its measurement were first used in Anglo-Saxon countries where the evaluation of health status has had a long history in public health and epidemiologic researches (6). The term HRQL has become preferred by many whose task is to assess the impact of disease and treatment on the lives of patients. HRQL focused on the impact of perceived health on an individual's ability to live a fulfilling life. Also, it takes into account levels of physical, mental, social, and role functioning, and includes abilities, relationships, perceptions, life satisfaction, and well-being. Moreover, HRQL is a concept that tries to embrace the spirit of the WHO definition of health by including both personal health status and social well-being when assessing health (7).

Due to prolonging the life expectancy, chronicle diseases such as cardiovascular, pulmonary or malignant, because of its importance, replace the acute, and the interest of medicine moves from simple extending the life, toward improvement of quality of life, what becomes the new model of modern medicine (3).

Quality of life and health-related quality of life - Definition

In everyday speech, QoL suggests impact of many external conditions and personal features. Thanks to that, an individual can feel satisfaction and dissatisfaction, he/she can plan keeping or changing the conditions in which lives. Public workers and media use the term related to the environment, physical and social: air pollution, soil and water pollution, living standards, and crime rates (8). The term is used in philosophy literature, sociology, geography, health economics, medicine, psychology, and pharmaceutics industry.

Quality of life is separated from the term health-related quality of life, so the consensus has been reached among experts on two important issues in the HRQoL field (8,9):

- it is recognized that the patient rather than a doctor or a nurse is the best source for obtaining HRQoL information.
- HRQoL is viewed as a multidimensional concept, which should include the four primary dimensions: physical functioning, encompassing self-care activities (eating, dressing), physical activities (walking, climbing stairs), and social activities (working, household, school); physical symptoms related to the disease or treatment (pain, diarrhea, neuropathy); psychological functioning, including emotional state and cognitive functioning; social functioning referred to the activities and association with friends, relatives and other social contacts.

There are many definitions for the QoL term, because of different approaches while considering it. Its meaning is differently explained and it depends on the user's age and position in social and political structure (10). QoL definition can be separated in general definitions, definitions specially related to health, and related to disease (11,12) (Table 1).

Table 1. General definitions and definitions specifically related to health and disease

AUTHOR	GLOBAL DEFINITIONS
Caiman, 1984	The extent to which hopes and ambitions are matched by experience
Ferrams and Powers, 1985	An individual's perceptions of well-being that stem from satisfaction or dissatisfaction with dimensions of life that are important to the individual
Grant et al, 1990	A personal statement of the positivity or negativity of attributes that characterizes one's life.
AUTHOR	**DEFINITIONS SPECIFICALLY RELATED TO HEALTH**
Schipper, 1990	A pragmatic, day to day, functional representation of a patient's physical, psychological, and social response to a disease and its treatment.
Cella and Tulsky, 1990	Patient's appraisal of and satisfaction with their current level of functioning as compared to what they perceive to be possible or ideal.
Gotay et al, 1992	A state of well-being which is a composite of two components: the ability to perform everyday activities which reflect physical, psychological and social well-being, and patient satisfaction with levels of functioning and the control of disease and/or treatment related symptoms.
WHOQOL Group, 1993	Quality of life is defined as an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychological state, level of independence, social relationships, and their relationships to salient features of their environment.
Testa and Simpson, 1996	The physical, psychological, and social domains of health, seen as distinct areas that are influenced by a person's experiences, beliefs, expectations, and perceptions.
AUTHOR	**DEFINITIONS SPECIFICALLY RELATED TO DISEASE**
Cella and Tulsky, 1990	HRQoL is more specific and more appropriate term than quality of life, because it refers to patients' assessment and satisfaction of their current level of functioning with it compared to what they consider to be possible or to be ideal.
Padila et al, 1998	The term HRQoL, connotes a personal, evaluative statement summarizing positive and negative attributes that characterize one's psychological, physical, social, and spiritual well-being at a point in time when health, illness and treatment conditions are relevant.

Measuring quality of life and health–related quality of life

Measuring QoL is important for making decisions especially about non-clinical aspects of disease, improvement of the doctor-patient relationship, in discovering of functional and psychological limitations, and in choosing the treatment in initial phase of disease, when the efficiency of applied therapy is temperate (for example drug just modify a disease). Also, it is important when you chose therapies that are little different, when you chose among a few efficient, different, clinical therapies, when there are dilemmas in applied therapies because of toxins, costs as well as for supplying information about using resources (13).

Measuring of quality of life and health related quality of life could be unidimensional and multidimensional (14). Unidimensional measuring refers to one dimension HRQoL. When they are used in clinical researches they can limit clinical information. They can show whether the treatment improves QoL, but they do not inform about the way of improvement. Multidimensional measuring is used in clinical researches. QoL assessment based on multidimension is important when there is a little information about the effects of a disease and/or treatment of a disease (15). Multidimensional measuring in the informal way points out which health intervention justifies invested money, but they cannot be used for cost benefits analysis (14).

Also, measuring QoL and HRQoL can be: global and specific (specific in relation to disease and in relation to medical treatment) (14). Global measuring is used in general population to measure health status of population and to compare different health conditions or diseases. They are also focused on the basic human values such as emotional well-being and on the possibility of everyday functioning (8,17,18). Specific measuring is related to the domains, which are important for a disease; and for different states, that has priority for a patient. Most usually they are used in clinical researches of drugs or therapeutics' intervention (8).

If HRQoL is included in the clinical research, it has three important characteristics. Firstly, the researchers, doctors, describe given conditions or a disease in terms, which are clinically important, and the patient can understand them easily. Secondly, it is that HRQoL domains can be independent predictor of the important clinical results - such as observing treatment, morbidity and mortality. These data insure precious consideration in history and prognoses of different states and diseases. Thirdly, we can get data about the

treatment, which determine individual daily functioning from the patient's point of view, what we can get or lose during the therapy. This can help a doctor to make decision to modify specific elements of therapy such as drugs, consultative health care, education of patients or service (19,20). All these information should be added to the information that the doctor gets during physical examination, laboratory tests and medical history. However, measuring HRQoL is used in small number of clinical researches as a primary goal, although quality of life is often better prognostic indicator than factors connected to the disease or treatment (14).

Measuring QoL and HRQoL can be done into three domains, that agrees with the health definition WHO: physical functioning (that includes symptoms, functional difficulties), psychological state (emotional and cognitive functions) and social interaction (work, daily activities, public relations). In case that measuring does not include one of these domains, HRQoL has negative assessment. However, the number of dimensions can be much higher (21).

Measuring quality of life – Instruments

The instruments for measuring quality of life can be global, generic and disease specific. Global measures (instruments) are designed to measure QoL in the most comprehensive or overall manner. This may be a single question that asks the respondent to rate his/her overall QoL or this may be an instrument such as the Flanagan Quality of Life Scale that asks people to rate their satisfaction in 15 domains of life (22). Generic measures (instruments) have much in common with global measures, but they are designed primarily for description. They are used in general population for the assessment of health status, different conditions or diseases. Usually, they are not specific for a particular disease or vulnerable population of patients and they are much more useful in general health researches, comparisons of different diseases and several studies. General instruments include large number of quality of life dimensions but in the first place physical, mental and social dimension (8,23). It is believed that the lack of the general instruments is their inability to identify condition - specific aspects of disease that are significant for the measurement QoL. If the data is necessary for major number of conditions, the instruments would have to be of enormous length. Additionally to specific instruments for a disease is needed to detect important clinical changes (22-24)

Disease specific instruments are orientated on the domains most relevant to the disease, condition or characteristics of patients. If the particular treatment or clinical trial is used, instruments could be called "treatment specific" or "trial specific", or by one name "situation-specific" (25). Specific instruments are needed for their homogeneity/brevity, and to ensure sensitivity for sometimes small, but clinically significant changes in health status and intensity of a disease. The recommendation is to use the combination of generic and specific instruments in the case when the overall QoL instruments are not satisfactory for specific diseases (26).

Cultural adaptation: The steps in the cultural adaptation

The cultural adaptation demands use of a proper language so that the translated questionnaire should be conceptually equivalent to the original and clear and understandable for a patient.

The process of cultural adaption could be done in cooperation with the Mapi Research Institute, Lyon, France. The cultural adaptation (translation) undergoes three steps: forward translation, backward translation and patient testing.

Forward translation

Forward translation consists of a few phases: engagement of two professional translators, making reconcile - the first intermediary version (forward translation), making the report for Mapi Research Institute and making the final first intermediary version.

The native language of the engaged professional translators must be their mother tongue and their English must be very good, too. They are independent in translating instructions for use and the questionnaire (instruction for filling, questions and responses) and they produce two versions of the forward translation (every translator gives an independent forward translation).

The reconciled - first intermediary version is created during the meeting of both translators and the local project manager. The translators compare their translations among themselves and compare them with the original questionnaire. The aim is to produce a conceptually equivalent translation of the original questionnaire and the language used, which should be colloquial and easily understandable.

The project manager makes the report for the Mapi Institute for each question in English. Also, the project manager explains translation problems, difficulties in translation, offers and accepts solutions and options of the first reconciled - intermediary version of forward translation, explaining translation problems, disagreements of the translators in the translation, offered and accepted solutions.

The final reconciled, intermediary version of the translation arises after the Mapi Institute has analyzed the report and after their suggestions have been added into the first intermediary version.

Backward translation

The backward translation implies a few phases: the engagement of the professional translator, making backward translation, loading the changes into the first intermediary version, making the report for the Mapi Institute and making the second intermediary version.

The native language of the engaged professional translator must be English and the language he/she translates must be very good, too. His/her task is to translate the first reconciled intermediary version of the forward translation into English as more literal as possible. The translator must not see the original English version of the questionnaire before he/she begins to translate.

Backward translation emphasizes disagreements and differences (that exist) between the first intermediary version and the original questionnaire. This is achieved by translating the backward translation and the original questionnaire. The aim of the meeting between project manager and translator is: to go carefully though the whole questionnaire, question by question, sentence by sentence and make comparison of three documents (the backward translation into English, the English original questionnaire and the first intermediary version for each single part of the questionnaire).

The differences that the project manager and the translator of the backward translated version should notice when making the comparison may include: faulty backward translation, faulty forward translation and structural differences between backward translation and the original questionnaire.

The revision of the whole questionnaire should be made at the meeting between the project manager and the translator. Also, project manager establishes the changes that should be made to the first intermediary version. The first intermediary version with the report of modification after back-translation, and the back-translation itself are sent to the Mapi Institute. The report should mention all the discrepancies between the back-translation and the English original as well as the explanations of all potential differences caused by faulty backward translation or faulty forward translation, or structural differences between backward translation and the original questionnaire. Also, the report should mention the explanation of the changes that have or have not been brought in the first intermediary version.

Mapi Institute reviews the backward translation and the report. All disagreements with respect to the original questionnaire are discussed with the local project manager. The second intermediary target version arises after agreement on all the changes that were made into the first intermediary version.

Patient testing, or cognitive debriefing

This step, "Patient Testing", includes: testing of the second intermediary version of the questionnaire, making reports for Mapi Institute, acceptance of the second intermediary version or making the third intermediary version that would be clearer than the previous one and more acceptable for all persons who use it. Mapi Institute should engage translators whose native language is local language and their task is to make the final version of the questionnaire.

The aim of the patient testing is: to test the comprehension and acceptability of the second intermediary version; to identify questions that are problematic as well as the reason for it; and to write down possible suggestions for understanding the formulation of questions.

The second intermediary version questionnaire is tested on a panel discussion, face to face with five patients who are suffering from the specific disease to which the questionnaire is referred. The idea is to choose five patients who would be representatives of the patients' population in a specific country. There are some criteria that are recommended while choosing patients including their education, profession and age.

When we speak about education, it is better that patients are with a lower level of education. Previous experiences have shown that people with a high level of education (professors, teachers, scientists, and doctors) never have difficulties in understanding while testing the questionnaire. It is preferable to have patients from several professional groups, but this should not be in contradiction with their education.

The role of the project manager is to discover all misunderstandings or misinterpretations and to identify words or wordings that may be inappropriate and to write them down. For the project manager is also important to express patient's feeling when answering some questions (face expression shows agreement or disagreement).

Throughout panel discussion, the project manager asks questions to the respondents about their general impression about the questionnaire: is it globally clear, easy to understand, easy to answer, is it too long, is it adapted for the condition, are the instructions clear?

After that, together with patients, the project manager goes through to whole questionnaire, question by question and checks the following:

- Are the questions difficult for understanding? If so, why?
- Are the offered answers clear and consequent with the questions?
- Is the primary concept of questions interpreted correctly? Is there any ambiguous formulation that would make more than one possible interpretation?
- Is the language used easy to understand, and is the language used as a common speech?

Next, the project manager makes one independent report of the panel discussion. He/she has to explain suggested changes that project manager finds to be relevant and the changes he/she suggests to be kept.

After the report has been examined and after discussion of patient testing results with the Mapi Institute, the third intermediary version of the questionnaire is compiled by integration of all changes into the next intermediary version. It is also possible to keep the second intermediary version if there are no significant changes.

Mapi Institute engages two local translators whose native language is the target language and their English is also very good, so they can translate the third (or, the second) intermediary version of the questionnaire into English. During the meeting of these two translators, they compare translations to the original version.

Changes that local translators suggest are discussed with the project manager. The final version of the questionnaire is created and is based on the results of this discussion.

Exercises

Measuring quality of life

The purpose of the exercises is to provide students with basic information about quality of life and measuring quality of life.

Task 1

Comparison of dimensions between generic and specific instruments

Students work individually. The students are given the generic questionnaires SF - 36, SF - 12, SF - 8 and specific questionnaires that are available from: http://www.sf-36.org/demos/SF-36.html. Students should notice the differences between these questionnaires and discuss about their dimensions. Some of the students will report what they understand from the comparison. Time: 90 min.

Task 2

Use the website page (http://www.sf-36.org/demos/SF-36.html) to fill in the "demo" online available questionnaires and look at your psychometrically-based physical component summary (PCS) score and the mental component summary (MCS) score. Compare your scores with the national standards (if applicable for your country). Time: 180 min.

References and Recommended Readings

1. Wood-Dauphinee S. Assessing quality of life in clinical research: from where have we come and where are we going? J Clinical Epidemiology 1999;52(4):355-63.
2. Revecki DA, Kline Leidy N. Questionnaire scaling: models and issues. In: Staquet MJ, Hays RD and Fayers PM, ed. Quality of Life Assessment in Clinical Trials. Methods and Practice, New York: Oxford University Press; 1998. p. 157-168.
3. Awad AG, Voruganti LNP, Heslegrave RJ. Measuring Quality of Life in Patients with Schizophrenia. Quality of Life Assessment. In: Mallarkey G ed. Quality of Life Assessment. Last advances in the measurement and application of quality of life in clinical studies, Chester England: Adis International Limited; 1998. p. 29-44.
4. Bech P. Measuring quality of life: The medical perspective. Nord J Psychiatry 1992;46:85-9.
5. Joyce D, Acquadro C, Haugh M. A Proposed Cochrane Health-Related Quality of Life Methods Working Group (QoL MWG). Quality of Life Newsletter 1998;20:3.
6. Anderson RT, Aaronson NK, Bullinger M, McBee WL. A Review of the Progress Towards Developing Health-Related-Quality-of-Life Instruments for International Clinical Studies and Outcomes Research. In: Mallarkey G ed. Quality of Life Assessment. Last advances in the measurement and application of quality of life in clinical studies, Chester England: Adis International Limited; 1998. p. 1-20.
7. Wood-Dauphinee S. Assessing Quality of Life in Clinical Research: From Where Have We Come and Where Are We Going? J Clinical Epidemiology 1999;52(4):355–363,
8. Testa MA, Simonson DC. Assessment of Quality of Life Outcomes. N Engl J Med 1996;334(13):835-840.
9. De Boer BJ, van Dam FSAM, Sprangers MAG. Health Related Quality of Life Evaluation in HIV Infected Patients. In: Mallarkey G ed. Quality of Life Assessment. Last advances in the measurement and application of quality of life in clinical studies, Chester England: Adis International Limited; 1998; p. 155-168.

10. WHOQOL Group. Measuring Quality of Life: The Development of the World Health Organization Quality of Life Instrument (WHOQOL), Geneva: WHO, 1993.
11. Camilleri-Brennan J, Steele RJS.Measurement of quality of Life in Surgery. J.R.Coll.Surg.Edinb. 1999;44(4):252-259. 11-12.
12. Patrick DL, Erickson P. Health status and health policy. New York: Oxford University Press, 1993.
13. Michael M, Tannock IF. Measuring health related quality of life in clinical trials that evaluate the role of chemotherapy in cancer treatment. Canadian Medical Association Journal 1998;158:1727-34.
14. Shumaker SA, Anderson RT, Czajkowski SM. Psychological test and scales. In: Spilker B, ed. Quality of life assessment in clinical trials. New York: Ravan Press, 1990:95-114.
15. Ware JE. Conceptualizing and Measuring Generic Health Outcomes. Cancer 1991;67:774-779.
16. Hyland ME. Quality-of-Life Measures as Providers of Information on Value-for-Money of Health Interventions: Comparison and Recommendations for Practice. In: Mallarkey G ed. Quality of Life Assessment. Last advances in the measurement and application of quality of life in clinical studies, Chester England: Adis International Limited; 1998; p. 21-33.
17. Ware JE. Conceptualizing and Measuring Generic Health Outcomes. Cancer 1991;67:774-779.
18. Eiser C, Tooke J. Quality of life in Type II Diabetes In: Mallarkey G ed. Quality of Life Assessment. Last advances in the measurement and application of quality of life in clinical studies, Chester England: Adis International Limited; 1998; p. 115-123.
19. Kaplan RM. Quality of life measurement. In: P. Karoly ed. Measurement Strategies in Health Psychology. New York: John Wiley, 1985.
20. Revecki DA, Kline Leidy N. Questionnaire scaling: models and issues. In: Staquet MJ, Hays RD and Fayers PM, ed. Quality of Life Assessment in Clinical Trials. Methods and Practice, New York: Oxford University Press; 1998. p. 157-168.
21. Kaplan RM. Quality of life measurement. In: P. Karoly ed. Measurement Strategies in Health Psychology. New York: John Wiley, 1985.
22. Hutchinson A, Fowlert P. Outcome measures for primary health care: What are the research priorities? British Journal of General Practice 1992;42:227-31.
23. Goligher JC. Judging the quality of life after surgical operations. Journal of Chronic Diseases 1987;40:631-633.
24. Guyatt GH, Bombardier C, Tugwell P. Measuring disease specific quality of life in clinical trials. Canadian Medical Association Journal 1986;134:895-9.
25. Osoba D. Guidelines for measuring health-related quality of life in clinical trials. In: Staquet MJ, Hays RD and Fayers PM, ed. Quality of Life Assessment in Clinical Trials. Methods and Practice, New York: Oxford University Press; 1998. p. 19-36.
26. Bowling A. Measuring disease: a review of disease-specific quality of life measurement scales. Buckingham Philadelphia: Open University Press, 1995.
27. van Geldorp MWA, Heuvelman HJ, Kappetein AP, Busschbach JJV, Takkenberg JJM., Bogers AJJC. The effect of aortic valve replacement on quality of life in symptomatic patients with severe aortic stenosis. Netherlands Heart Journal 2013;21(1):28-35.
28. Nauck T, Lohrer H. Translation, cross-cultural adaption and validation of the German version of the Foot and Ankle Ability Measure for patients with chronic ankle instability. British Journal of Sports Medicine 2011;45(10):785-90.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Distance learning in health education**
Module: 2.27	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Izet Masic, MD, PhD, Professor Chair for Medical Informatics, Medical faculty, University of Sarajevo, Bosnia and Herzegovina
Address for correspondence	Izet Masic, MD, PhD, Professor Chair for Medical Informatics, Medical faculty, University of Sarajevo, Bosnia and Herzegovina www.imasic.org; E-mail: imasic@lol.ba
Keywords	Bosnia and Herzegovina, distance learning.
Learning objectives	After completing this module students should: • know the definition and characteristics of tele-education; • be familiar with advantages and disadvantages of tele-education in comparison to classical methods of education.
Abstract	Increase and development of distance learning (DL) technologies over the past decade has exposed the potential and the efficiency of new technologies. Benefit and use of contemporary information technologies is the area where medical informatics got the most on understanding and importance. Definition of distance learning as "use of technologies based on health care delivered on distance" covers areas such as electronic health, tele-health (e-health), telematics, telemedicine, tele-education, etc. as web-based education. For the need of e-health, telemedicine, tele-education, web-based education and distance learning there are various technologies and communication systems from standard telephone lines to the system of transmission digitalized signals with modem, optical fiber, satellite links, wireless technologies, wedsites, etc. Web-based education represents health education on distance, using IC technologies and Internet, as well as continuous education of a health system beneficiaries and use of electronic libraries, databases, or electronic data with databases of knowledge. The number of events on distance learning and web-based education organized by the teaching staff from the Chair for Medical Informatics, Medical Faculty, University of Sarajevo is very broad including: professional and scientific events, workshops and congresses, tele-exams, introduction of distance learning in the curriculum of biomedical faculties, etc.
Teaching methods	An introductory lecture gives the students a first insight into the characteristics of distance education. The theoretical knowledge is illustrated by a case study. After introductory lectures, students first carefully read the recommended readings. Afterwards, students discuss the characteristics of distance education with other students.
Specific recommendations for teachers	• ECTS: 0.2 • work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a computer room; • equipment: computers (one computer for 2-3 students), LCD projection equipment, internet connection, access to the bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire.

DISTANCE LEARNING IN HEALTH EDUCATION

Izet Masic

Theoretical background

Distance learning - definition and description

"There are two great equalisers in life - the Internet and education. By combining the two, e-learning will be the great equaliser in the next century. By eliminating barriers of time, distance, and socio-*economic status, individuals can now take charge of their own lifelong learning."*

John Chambers, CEO of CISCO systems

Distance learning is conventionally defined as: any educational or learning process or system in which the teacher and instructor are separated geographically or in time from his or her students; or in which students are separated from other students or educational resources (1-3). The most important factor which influences the changes occurring in education has been the installation and development of the Internet and electronic multimedia techniques. Distance learning does not preclude traditional learning processes; frequently it is used in conjunction with in-person classroom or professional training procedures and practices.

Distance learning is used for self-education, tests, services and for the examinations in medicine, i.e. in terms of self-education and individual examination services. The possibility to work in the exercise mode will image files and questions is an attractive way for self-education (4-7). The standard format of the notation files enables to elaborate the results by commercial statistic packets in order to estimate the scale of answers and to find correlation between the obtained results. The method of multi-criterion grading excludes unlimited mutual compensation of the criteria, differentiates the importance of particular courses and introduces the quality criteria. By using computers and teleconferencing technology and through partnerships with local communities, institutions and the private sector, an open, effective, virtual learning community is now in place. Sites are located in college and university campuses, hospitals, schools, libraries, community centres and private companies. Courses are also being delivered to private homes.

For the need of e-health, telemedicine, and tele-education there are various technologies and communication systems from the standard telephone lines to the system of transmission digitalized signals with modem, optical fibber, satellite links, wireless technologies, etc. There is no doubt that Internet causes "revolution" in all of the above, and the latest possibilities are distribution of virtual medical instruments and medical data in real time and possibility of use in primary health care, even for some diseases with poor prognosis. This revolution on how information is stored, transmitted and accessed has extremely important implications for the health sector, especially now when embarking on a global effort to renew the tenets of Health for All based on primary health care and disease prevention, health promotion and costumer education, in the context of service delivery guided by the equity, quality, effectiveness and efficiency. According to Grimson et al. in Dublin, *"the need to participate in continuing professional development or* continuing medical education is considered to be at the very least highly desirable and more likely mandatory. The use of Information Communication Technologies (ICT) (Figure 1) is one way by which this can be facilitated in a timely and cost-*effective manner"* (7-10).

Traditional way of learning, and learning from the distance

The latest research shows that the format of instructions itself has no important influence on the students' achievements if access and availability to information technologies is assured as well as usage of the adequate content of education. In the assessment of the authentic situation the following issues should be addressed:

- results of different tests prepared by lectors has trend to show advantages in comparison with traditional learning methods and there is significant distinction in affirmative attitude to educational materials between distance and traditional learning;
- traditional methods demonstrate better organization and they are clearer in respect to distance learning;
- organization and needs for more efficient influence of distance learning very often improve traditional methods by teachers;
- future research should be focused on critical factors in determining student involvement in development of the educational process;
- the variety of teaching and learning options provided by technology allows education to be provided in an appropriate manner to a broader student demographic profile than ever before.

Figure 1. http://svajunas94.blogspot.com/

Facts about distance learning and tele-education

Distance learning enables permanent learning (lifelong learning), students can improve themselves professionally and independently, at their own tempo, at place and time that they choose by themselves, they can choose a great deal of subjects which offer different institutions, teachers-individuals; students go through materials for learning by speed of their own and as many times as they want. The place can be chosen – it depends on the medium which is used for learning materials (they can learn at work, or from home). Themes which are not offered by studies in that field – students find and attend the programs which they are interested in, although they are not offered by educational or business institutions in place in countries where they live in, or work. Taking part in top-quality and most prestigious programs – student can "attend" at least some studies at the top-quality institutions, or studies held by lecturers that are very famous experts without changing their place of living. Choosing this way of learning – active or passive learning, different kinds of interaction include: "classical" written material and writing down their own lecture notes, interactive simulations, discussion with other students (e-mail, tele-conferences).

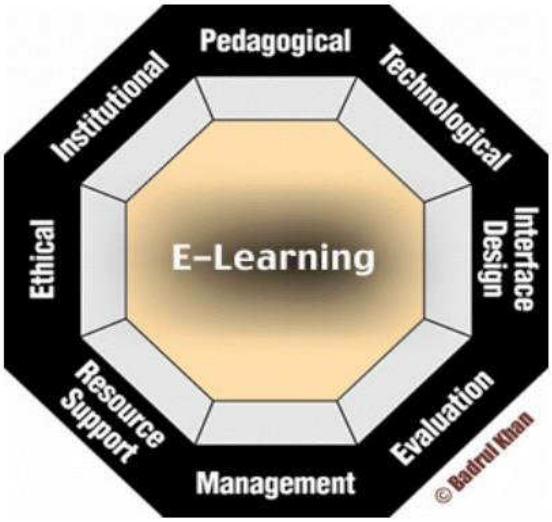
Practical work with different technologies – students get not just information about what they learn, but additional knowledge and skills about using computer, CD players and video recorders. Independent learning – teachers learn too from students who independently ask for information sources.

The meaning of education/learning in distance can be expressed by the definition that: it is a form of education which is in a process permanently, or most of the time, all or most of the tasks of teaching and learning separately during the time and space between teacher and student.

Pedagogical and organizational improvements have fundamental importance. It is in use both the interaction "teacher-student" and the interaction "student-student". Phases of synchronized learning are combined. Individual and group works are also combined. If all these forms are involved in the educational process, they mutually supplement each other, as a last resort. Traditional education as well as contemporary education is supported by informatics' technologies in a unique system of flexible education. In order to use advantages of flexible education, it is necessary to combine different forms of learning, during the preparation phase and development of every educational course in an appropriate way.

Distance learning is not simply a set of "infrastructure", but rather a concept of learning that incorporates different technologies and learning media. Within a province, different video, audio and computer tele-conferencing systems, along with Computer Based Training, Computer Managed Instructional systems and other media are being integrated technologically, instructionally and organizationally. The tele-education concept crosses all jurisdictions among institutions both within and outside a province, public and private, at any level of education, to anywhere including institutions, workplaces and the home. Tele-education, tele-teaching, tele-training, tele-mentoring, and tele-accreditation have been clearly demonstrated and are currently a common practice (Figure 2).

Figure 2. http://esladaptationreadings.blogspot.com/2011/04/e-learning-framework-considerations.html



Advantages and disadvantages of distance learning

Distance learning, compared to the traditional way of learning, has many advantages, as well as dis-advantages.

1. Advantages of distance learning

Some of the main advantages of distance learning are as follows:

- The economical factor (reduces the travel time and cost for attending face-to-face instruction);
- Student has 24-hour access to the required/needed information;
- He/she is given the opportunity to learn the subject in his/hers own time and speed;
- He/she can access learning material independently of place and time;
- He/she is given the opportunity to learn how to work independently;
- Using e-mail or chat rooms he/she is able to contact the professor or his/her assistant if there are any questions or confusions regarding the lectures;
- Online learning may be more student-centered if it allows students the option to select learning materials that meet their level of knowledge and interest;
- Proof of completion and certification, essential elements of training initiatives, can be automated (2,3,21).

Fundamental advantages of flexible education in terms of classical education are as follows:

- More efficiency;
- Increase capacities of educational institutions;
- Education can be easily adopted to the needs of education on-the-job;
- Costs of the educational process are smaller;
- It is possible to distribute the education uniformly, the new educational programs are available for fields outside the educational and economic centres;
- It enables the possibility of access to the foreign educational resources to the various institutions;
- Superior quality of the knowledge gained.

2. Disadvantages of distance learning

Many critics consider that using e-mail or chat rooms to obtain contact with the professor is actually the main disadvantage of this system of learning. Question arises whether this way of professor-student communication is helpful to students because face-to-face contact is missing as well as the opportunity of student-professor relationship building. Disadvantages to the trainer or organization include the following:

- Up-front investment required of an e-learning solution is larger due to development costs. Budgets and cash flows will need to be negotiated;
- Technology issues that play a factor include whether the existing technology infrastructure can accomplish the training goals, whether additional “tech” expenditures can be justified, and whether compatibility of all software and hardware can be achieved;

- Inappropriate content for e-learning may exist according to some experts, though limited in number. Even the acquisition of skills that involve complex physical/motor or emotional components (for example, juggling or mediation) can be augmented with e-learning;
- Cultural acceptance is an issue in organizations where student demographics and psychographics may predispose them against using computers at all, let alone for e-learning (21).

Also, there are some disadvantages for the Learner:

- Technology issues of the learners are most commonly technophobia and unavailability of required technologies;
- Portability of training has become a strength of e-learning with the proliferation of network linking points, notebook computers, PDAs, and mobile phones, but still does not rival that of printed workbooks or reference material;
- Reduced social and cultural interaction can be a drawback. The impersonality, suppression of communication mechanisms such as body language, and elimination of peer-to-peer learning that are part of this potential disadvantage are lessening with advances in communications technologies (21).

Fundamental disadvantages are as follows:

- Learners with low motivation or bad study habits may fall behind;
- Without the routine structures of a traditional class, students may get lost or confused about course activities and deadlines;
- Students may feel isolated from the instructor and classmates;
- Instructor may not always be available when students are studying or need help;
- Slow Internet connections or older computers may make accessing course materials frustrating;
- Managing computer files and online learning software can sometimes seem complex for students with a beginner-level computer skills;
- " Hands-on", or lab work is difficult to simulate in a virtual classroom (22).

Case study

Distance learning in medical curriculum and implementation of distance learning at Medical Faculty, University of Sarajevo, Bosnia and Herzegovina

Background

By late 1994, the Internet included 3.2 million computer nodes spread across more than 57,000 institutions in more then 80 countries, with an estimated 30 million users. By the end of the century the Internet linked more than 400 million persons.

The 2002 Eurobarometer survey showed that an average of 78% of EU medical GPs were online, with (at the highest level) 98% in Sweden and 97% in the United Kingdom. The number of "online patients" grows every day as well. The 2003 Eurobarometer survey on health information sources showed that 23% of Europeans used Internet for health information and that 41% of the European population considered that Internet is a good source of information on health (3).

In spite of the fact that Bosnia and Herzegovina (B&H) was the last (or, second before the last) country in Europe in the use of Internet technologies, there is a group of enthusiastic people accompanying Prof. Izet Masic, MD, at Medical Faculty and University of Sarajevo who have been making significant efforts to improve the poor digital literacy in the University and among medical professionals. E.g., every 60^{th} citizen of B&H uses Internet (50.000 citizens of 3.000.000 citizens in B&H), while in Slovenia every forth Slovenian is familiar and uses Internet technologies, for example.

In October 2003, the University of Sarajevo started the distance learning education by opening the University Distance Learning Centre. Opening of the University Distance Learning Centre, as a coordination body and leader in all activities in connection with distance learning, has provided an opportunity for development and growth of this kind of lifelong education.

In relation to the above-mentioned project conducted by the University Tele-Information Centre (UTIC) and as continuation of a two-year project on possibilities of introduction of distance learning in medical curriculum, the Cantonal Ministry approved and supported a new project namely: Introduction and Implementation of Distance Learning in Medicine. Platform for the course of distance learning is achieved in collaboration with UTIC. University Tele-Information Centre was established as part of the University of Sarajevo and first ISP in B&H in 1996 (UTIC). It is a scientific-organizational unit of the University of Sarajevo for improvement of scientific-research work and through UTIC members of the University can be gathered in the unique computer-communication structure (Figure 3).

Objectives of UTIC are as follows:

- to connect members of the University of Sarajevo with similar institution in the country and abroad due to more efficient use of scientific, research and educational resources;

- use of educational databases and other information for the needs of the University and its members;
- development an integration of informatics (computer technologies) in education;
- creation of flexible infrastructure which will enable e-learning to be accessible to all students and university staff;
- improvement of general digital literacy of academic population;
- development of top quality educational content which could be integrated in the actual European processes of the e-learning "revolution".

With the help centre for distance learning, "Lucis Centrum", was established (Figure 4). We hope that this is just a beginning step towards improvements of the B&H education system and that this project will serve as an indicator towards that future.

On the UTIC website, seven students (enrolled from the Medical Faculty, for the subject "Medical Informatics") are able to learn from a distant location. So far, teaching staff uploaded eleven lectures at the website (Figure 5):

1. Hardware and software
2. Medical documentations
3. Medical informatics
4. Methods of data manipulation
5. Nomenclatures and classification systems
6. Data organization
7. Data, information and knowledge
8. Lectures 1
9. System and communication
10. Structure and data organization, and
11. Expert systems

Figure 3. http://web.utic.ba/

Figure 4. Uploaded materials for the subject "Medical Informatics" at the webpage of Lucis Centrum of the University of Sarajevo

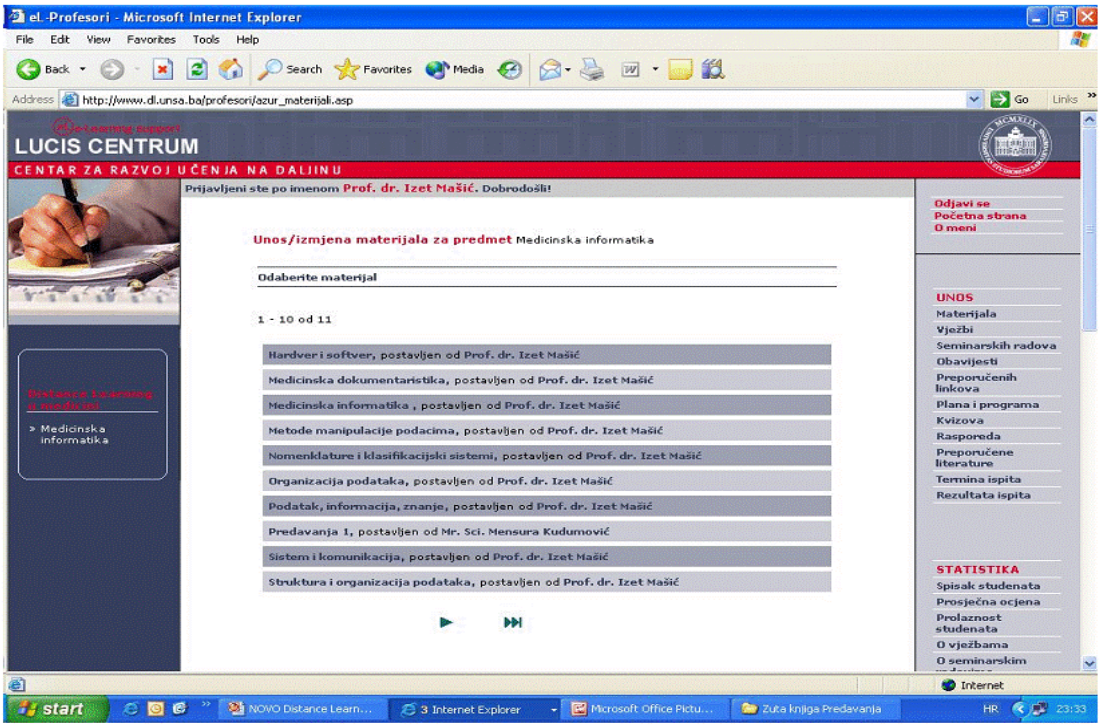
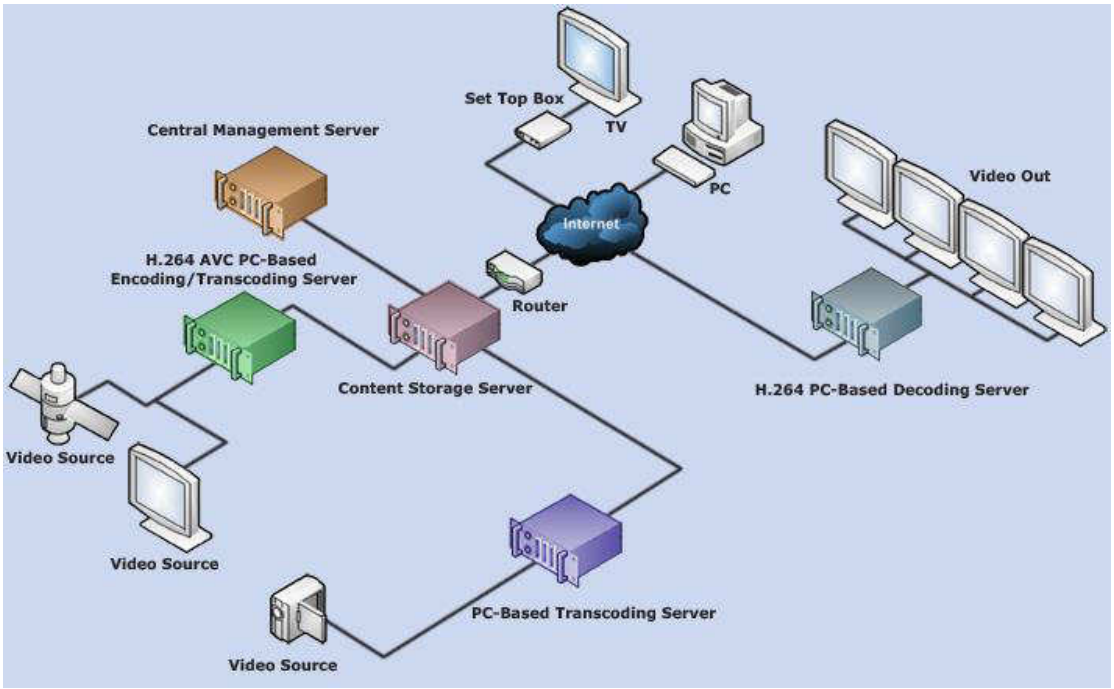


Figure 5. http://davispa.wordpress.com/2010/09/08/defining-distance-learning/



Beside the materials, it is possible to upload and download the following:

- practical works,
- seminar work,
- information,
- recommended links,
- plan and programs,
- quiz,

- schedule,
- recommended readings,
- examination schedule, and
- examination results.

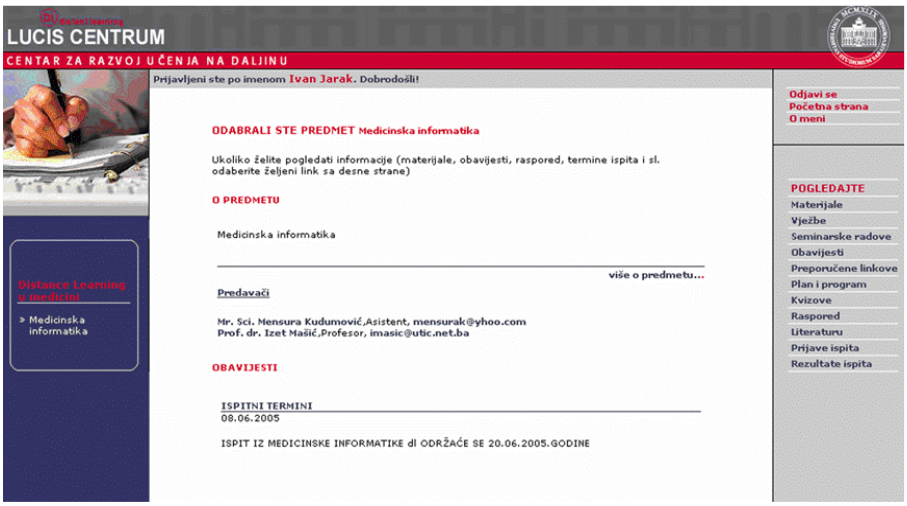
Basically, software application has two interfaces: teacher and student interface. Access from any of these is very simple and fast (1,2,3).

Education content of distance learning

Lecture contents will be presented in our virtual classroom. In our case, learning material from the subject of medical informatics, and later, hopefully from other medical subjects, will be available on web site, www.e-learning.ba (Figure 6).

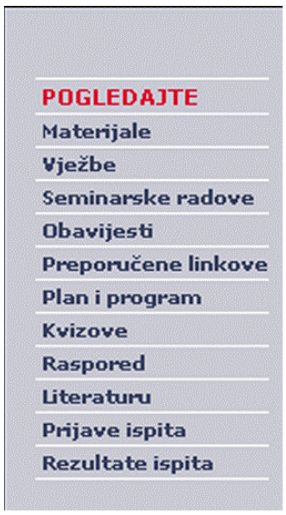
In this "classroom", learning materials, power point lecture presentations as well as practice exercises with step-by-step instructions, are easily accessible to the students. Moreover, on this website, students will be able to find subject relating literature as well as English version of the presentations (Figure 6). To access this information requires only one click in the download option (Figure 7), as well as one second of patience depending on student's Internet speed connection. In short, our virtual classroom gives students the opportunity to access the required information at any time, and in any place without having to be bound to the classroom.

Figure 6. Content of Medical Informatics page at the webpage of Lucis Centrum of University of Sarajevo



Students are able to browse through the "classroom" using standard navigations (Figure 7):

Figure 7. Navigator bar at the webpage of Lucis Centrum of the University of Sarajevo



Using these simple navigations, maximal efficiency and fast access to needed materials, is possible.

As we can see in Figure 7, all links are in chronological order according to the plan and program of the lecture as well as practices (materials, practical works, seminar work, information, recommended links, plan and programs, quiz, schedule, recommended readings, registration to examination, and examination results).

Figure 8. An example of a "Quiz" window at the webpage of Lucis Centrum of the University of Sarajevo

Special attention is given to the link "kvizovi" (quiz). In order for the student to check his/hers progress (to test his/her knowledge of the lecture he/she studied), every lecture is followed by quizzes. Quiz questions are multiple-choice questions (Figure 8) and they are based on the lecture content. After every quiz, the student receives "feedback" regarding his/her progress. Results are given in terms of percentages (one needs 51% right answers to pass the quiz). In this way, the student has absolutely control over his/her work.

Advantages and disadvantages of distance learning in Bosnia and Herzegovina

In many universities across B&H, students' contact with professors are almost impossible (unless one needs to orally take the exam), due to many other jobs or responsibilities of the professors; hence, students are mainly able to communicate with professor's assistants. Moreover, through the traditional way of teaching, during the lectures, students obtain mostly the information which they can find in the literature or in the Internet. Rarely, there is a student-professor interaction, or a lecture discussion during the class. From this, one can conclude that an ambitious student using tele-education will experience minimum lose.

We live in the "Age of Information". These technologies are changing the way we perceive the world, how we think and communicate with each-other. Established cultures are being transformed and new cultures are being formed. New virtual environment affects the way we build our sense of who we are.

Some characteristics of the Internet which people in B&H should have some basic knowledge include the following:

- large volume of users and potential users;
- lack of physical boundaries which allows for the manipulation of time and space;
- information can be accessed in a concurrent fashion using different media;
- concept of redundancy.

In the virtual environment we are applying for information in a way that is expanding our senses and one must to take into account that experience is occurring in the context of the virtual environment. Information without a context has no meaning.

Our experience in the application of tele-education at biomedical universities

The greatest progress was made in the area of tele-education and distance learning in B&H. Distance learning does not preclude traditional learning processes; frequently it is used in conjunction with in-person classroom or professional training procedures and practices. Distance learning is used

for self-education, tests, services and for the examinations in medicine, i.e. in terms of self-education and individual examination services. The possibility to work in the exercise mode with image files and questions is an attractive way for self-education (12-15).

The very first serious initiative was generated by the World University Service of Austria (WUS Austria) in B&H. During 2002 and 2003, WUS Austria, through its programs, distance learning 2002 and distance learning 2003, supported the development of the educational processes at B&H universities. Near the Chair for Medical informatics at the Medical Faculty of the University of Sarajevo, since 2002 there is in progress a project named: "Possibilities of introducing distance learning in medical curriculum", approved by the Federal and the Cantonal Ministry of Science and Education. The purpose of the project is to facilitate improvement of educational process at biomedical faculties, applying contemporary educational methods, methodologies and information technologies in accordance with the strategy and goals proclaimed by Bologna Declaration. A pilot project was realized during school years 2003-2005, and theoretical and practical education of the subject "Medical Informatics" are adapted to the new concepts of education using world trends of education from the distance. One group of students was included in the project finalized by an electronic exam registration and an electronic exam held publicly on 20th June 2005, in the Physiology amphitheatre of the Medical Faculty in Sarajevo (Figure 9,10).

Figure 9. First electronic exam held publicly on 20th June 2005, in the Physiology amphitheatre of the Medical Faculty in Sarajevo, Bosnia and Herzegovina

Bologna process, which started across European countries, provides us to promote and introduce modern educational methods of education at biomedical faculties in B&H. The Chair of Medical informatics and the Chair of Family Medicine at the Medical Faculty of the University of Sarajevo started to use web-based education as a common way of teaching of medical students. Satisfaction level with this method of education within the students is quite high, but not yet suitable for most of medical disciplines at biomedical faculties in B&H. Websites of the Chair for Medical Informatics and Family Medicine and UTIC are shown in Figure 11.

Bosnian Society for Medical Informatics (BHSMI) is very proactive in promoting telemedicine and tele-education as part of it. The last event organized by BHSMI consisted of a Special Topic Conference named e-Health and e-Education, which was held in December 2005 (Figure 12).

Figure 10. First electronic exam held publicly on 20th June 2005, in the Physiology amphitheatre of the Medical Faculty in Sarajevo, Bosnia and Herzegovina

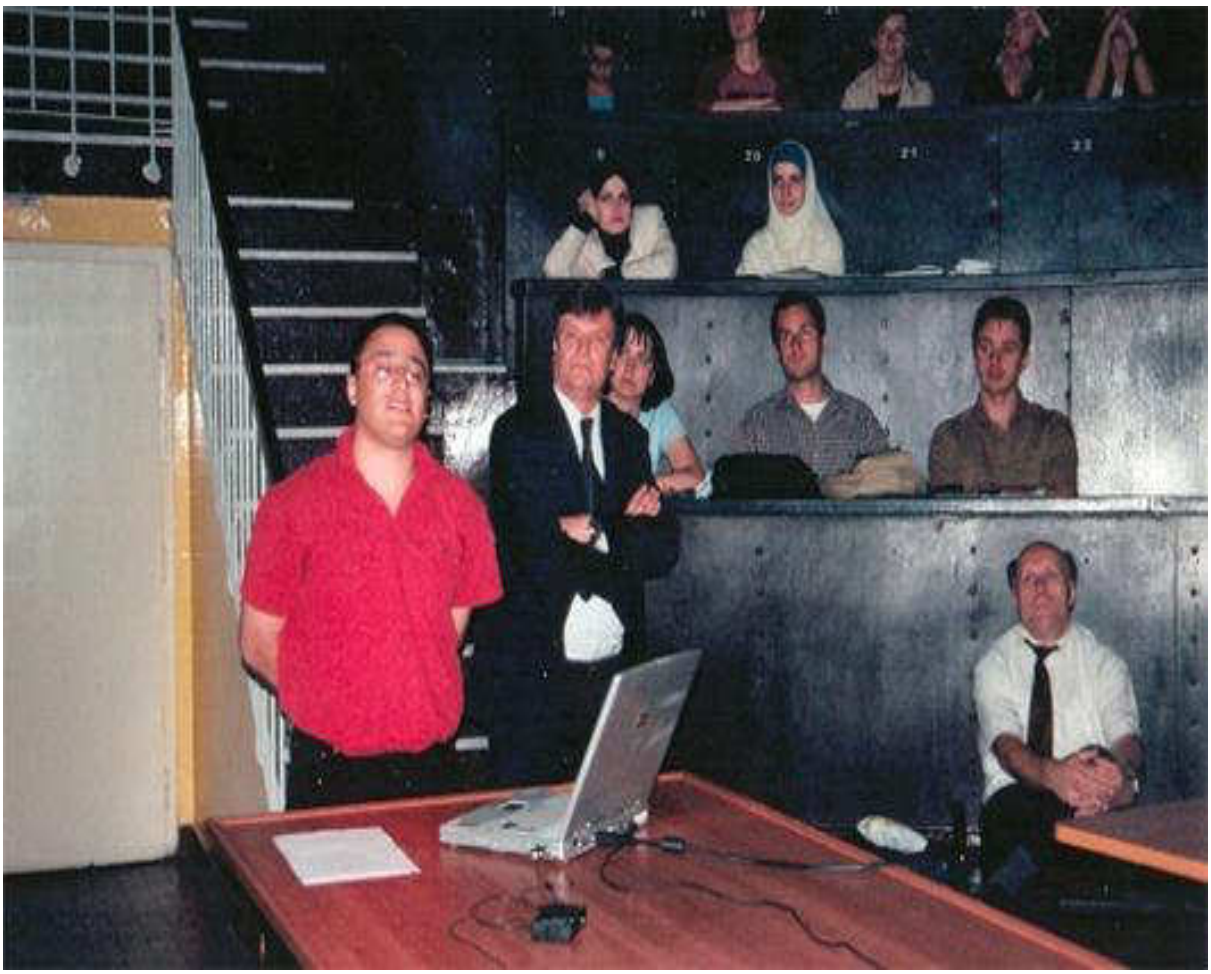
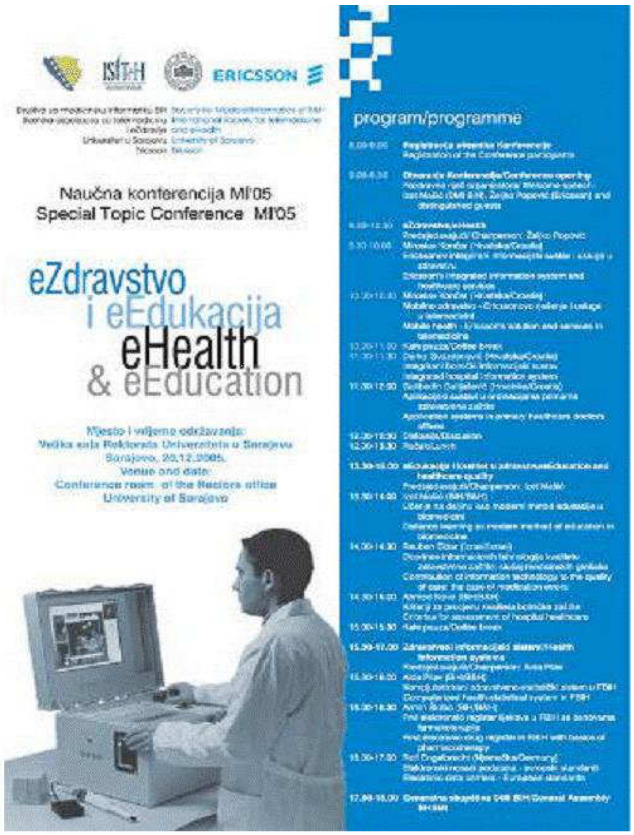


Figure 11. Websites of the Chair for Medical Informatics, the University Tele-Information Centre of University of Sarajevo, and the Chair for Family Medicine at Medical faculty of the University of Sarajevo

Figure 12. Special topic Conference: e-Health and e-Education



Distance learning in medicine has impact on telemedicine and practicing medicine as well. Basic skills of the use of computers and networks must be a part of all future medical curricula. The impact of technical equipment between the patient and the doctor must be understood, and the situation where the diagnosis based on live voice or picture is different from a normal doctor-patient contact. In some areas, telemedicine requires unique techniques. Tele-robotical disciplines differ from what surgeons normally learn. Telemedicine, and distance learning as a prerequisite, is les suited for doctor-to-doctor consultation, and the first contact to a doctor should always be a face-to-face consultation (15-19).

Where are we now?

Virtual environments are getting more and more popular: we are talking about virtual libraries, virtual schools, virtual universities, and virtual laboratories (IITAP & UNESCO 1999) to broaden access to knowledge and skills as far as connectivity is available. In combination with traditional face-to-face classroom-based teaching and learning, virtual environments become efficient means in translating the principle of the right to access to knowledge into reality. ICT can now offer effective and efficient learning environments for primary and secondary education, vocational training, universities, continuous education, but also for literacy training (23,24).

Today, many schools are introducing a system of distance learning, allowing students' access to educational materials. In the past five years, there have been made a lot of interactive virtual classrooms for medical students and other faculty, too. Mass published a new online atlas from all fields of medicine that help students learn in a 3-D form of medical items.

The message from the American Society for Training and Development conference last year was very clear: E-Learning is the future and the challenge is set for everyone to get on board. The World Wide Web is expanding at a tremendous rate and through the Internet we are all going to realize the huge benefits of E-Learning.

In the USA, they predicted that the e-learning market was going to grow from the present 7% to 60% in just three years. The potential was limitless and almost all training needs could be met by e-learning in the future.

A year later, however, the situation is not quite as rosy, the reality is that e-learning is finding its level and not everything can be done over the Internet.

Clearly, the opportunities to practice newly acquired skills and receive trainer feedback, are best suited to the traditional classroom environment. However, the opportunities for learners to interact with

each-other and participate in the activities of the learning process can now be achieved very effectively using the Internet. Facilities such as discussion forums, chat rooms and email are all widely available on the internet and can facilitate collaboration between learners and their trainers. With these types of facilities we can now extend the learning process beyond the classroom and into the workplace.

Figure 13. http://www.efsa.unsa.ba/dl/

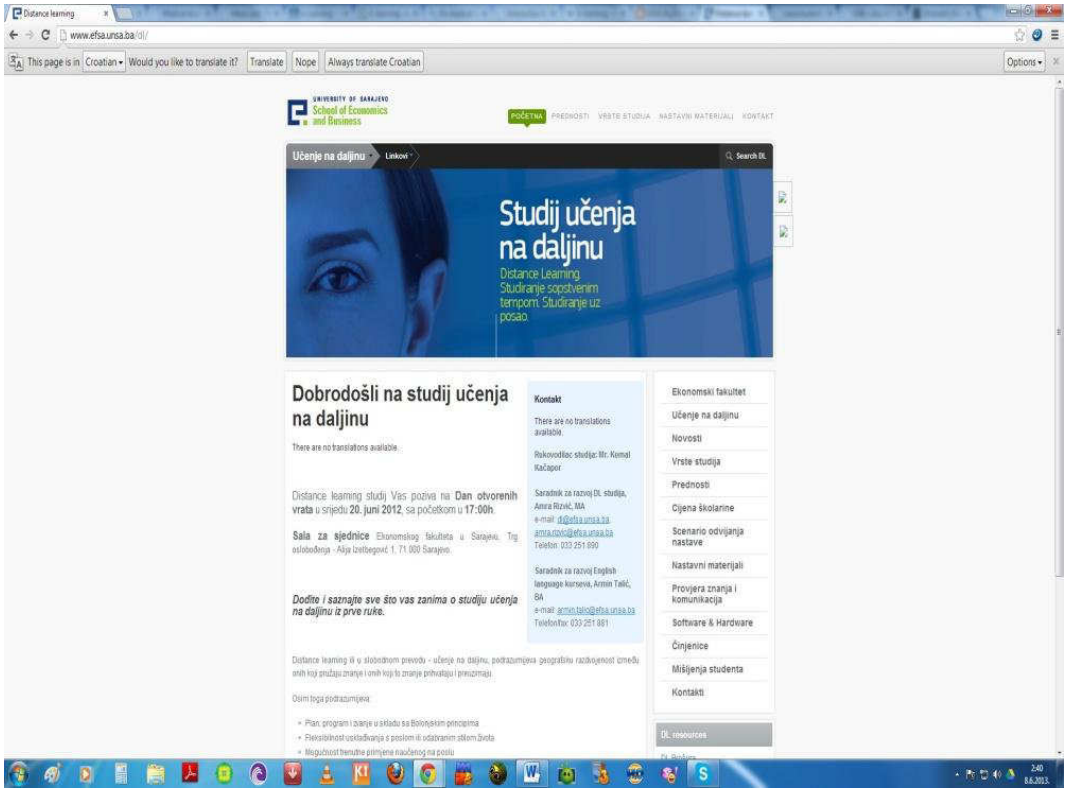


Figure 14. http://www.fpn.unsa.ba/ba/

Figure 15. http://fpn.unsa.ba/enastava/

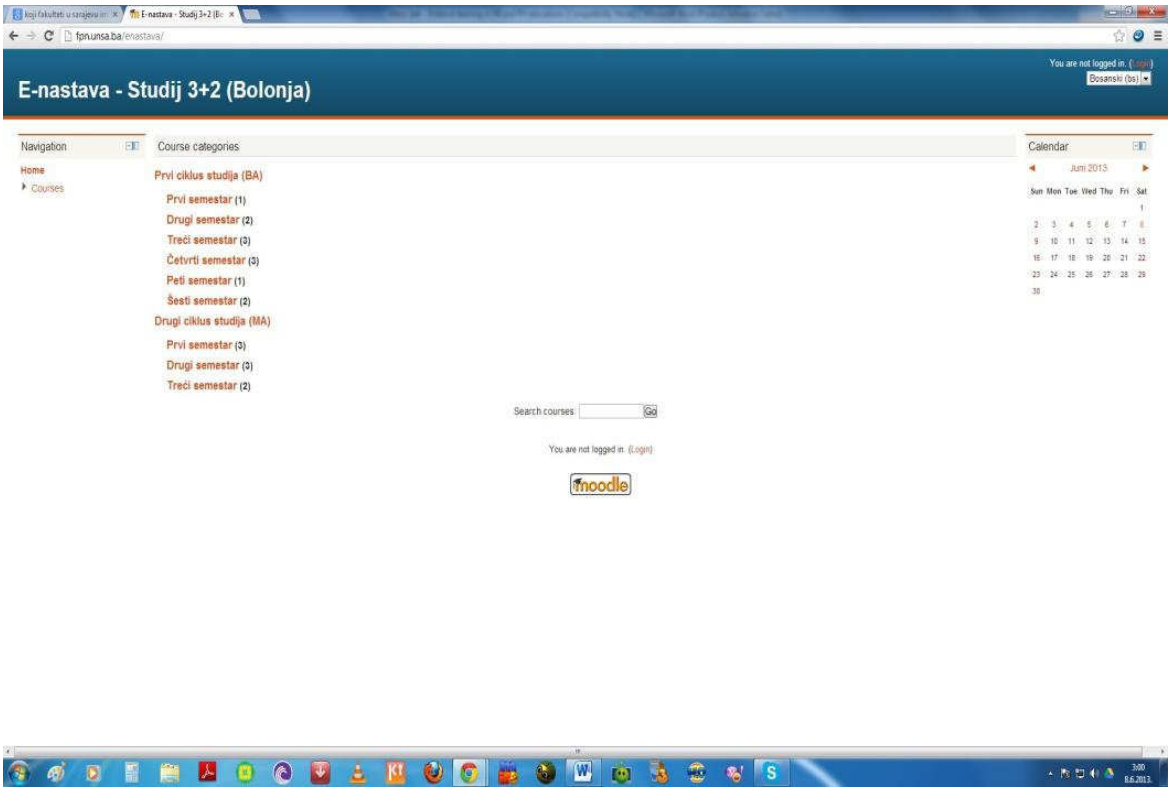


Figure 16. http://www.rcplondon.ac.uk/cpd/manage-your-cpd/cpd-approval-distance-learning

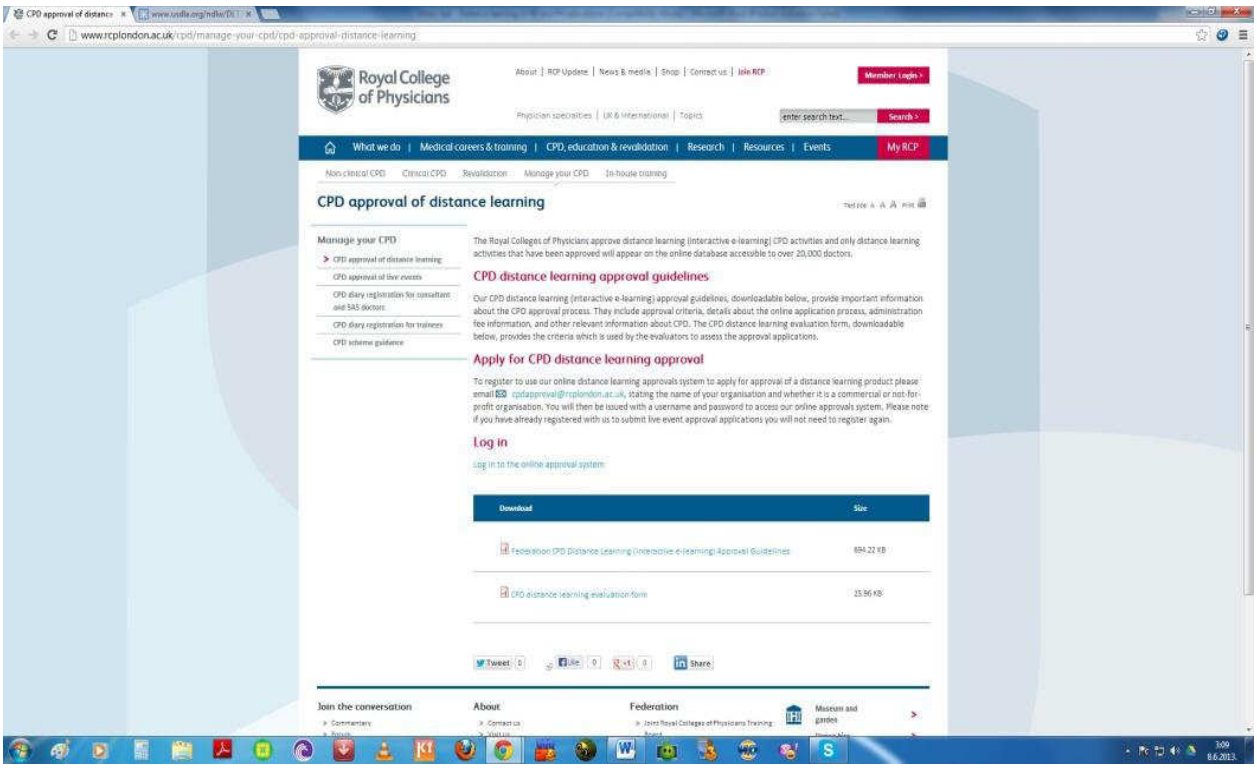


Figure 17. https://www.som.soton.ac.uk/learn/elearning/materials/3d/

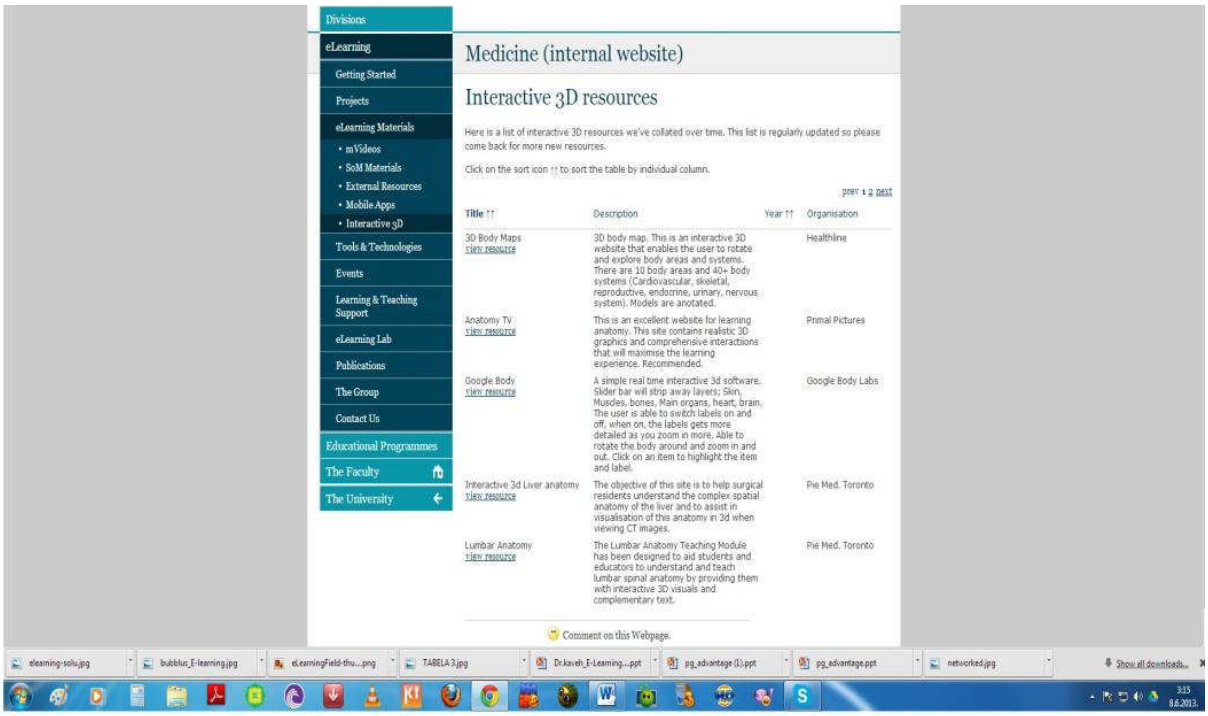


Figure 18. http://www.ztm.hr/

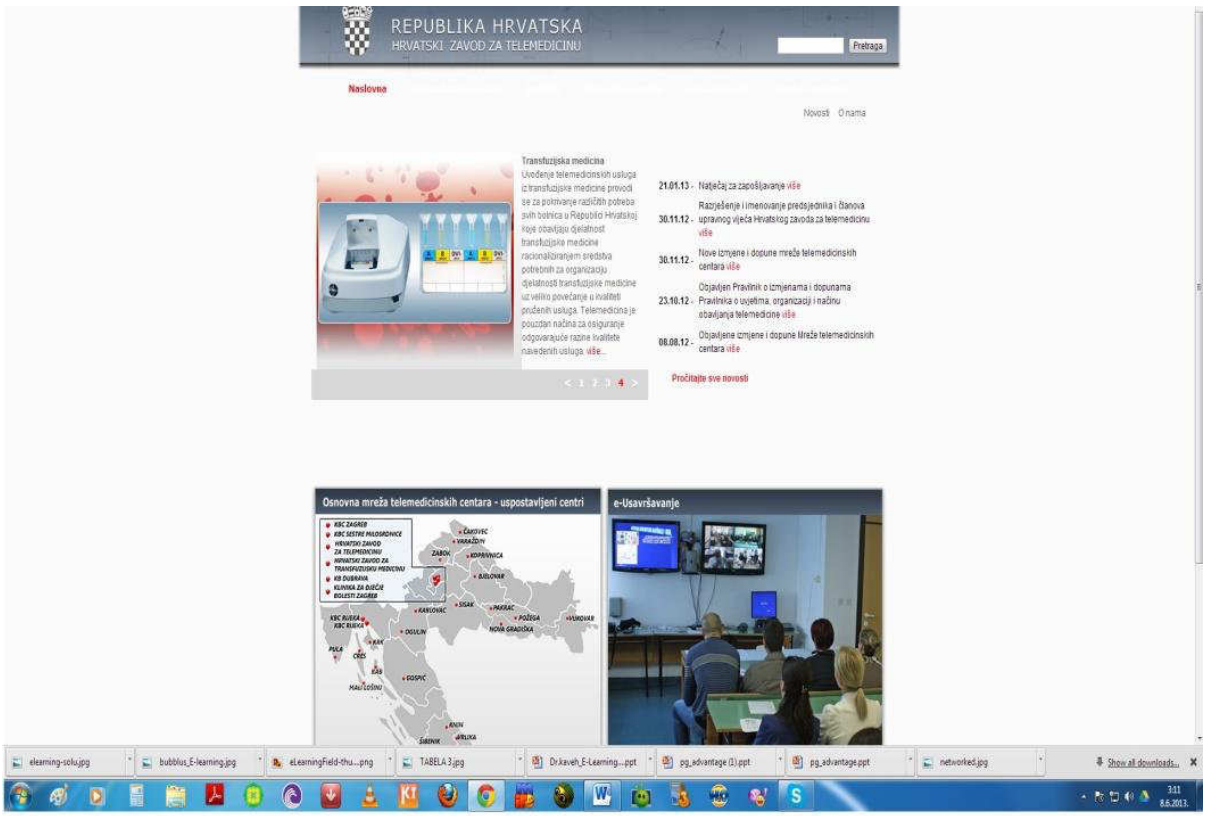


Figure 19. http://www.hawaii.edu/dl/

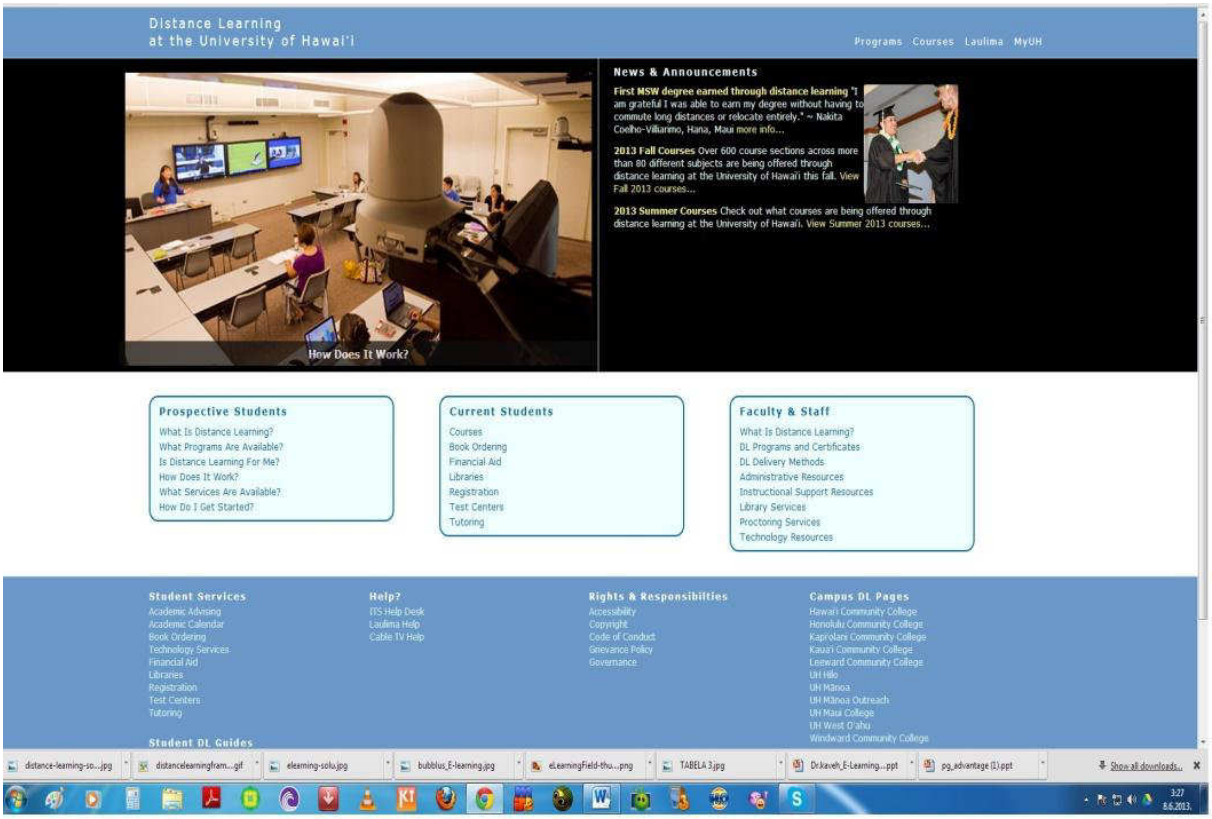
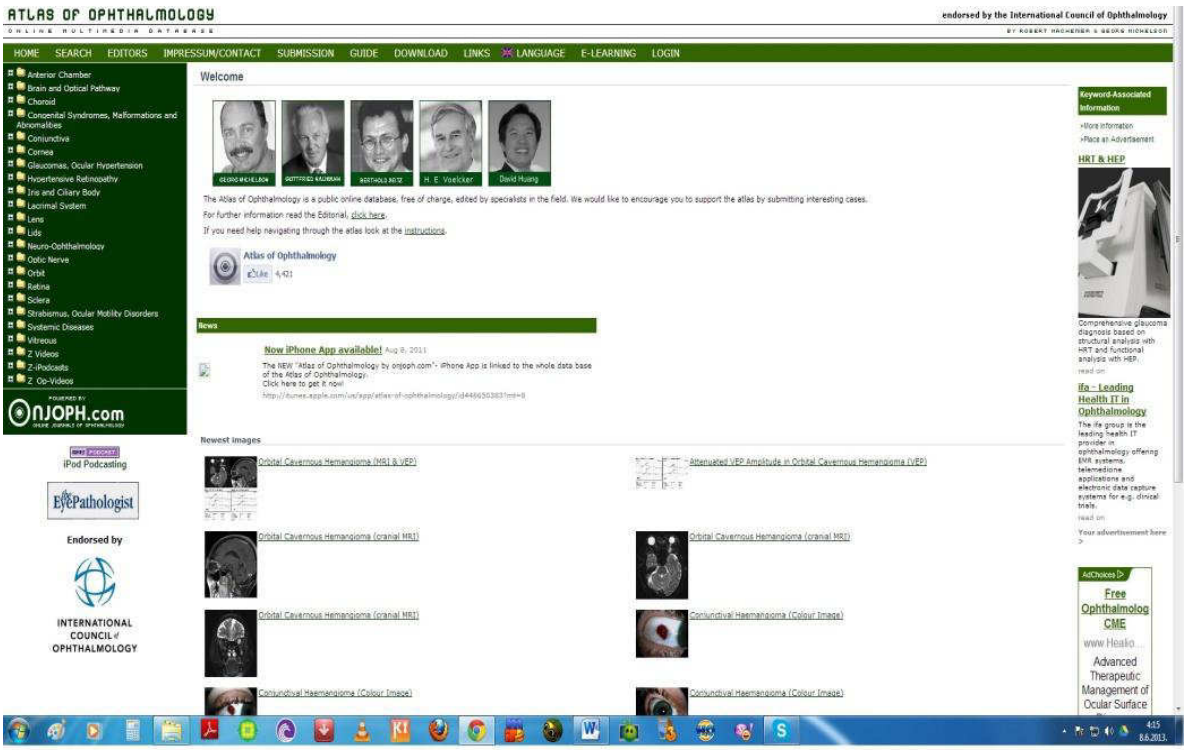


Figure 20. http://www.atlasophthalmology.com/atlas/frontpage.jsf?locale=en



Conclusions

Distance learning in medicine has impact on telemedicine and practicing medicine as well. Basic skills of the use of computers and networks must be a part of all future medical curricula. The impact of technical equipment between the patient and the doctor must be understood, and the situation where the diagnosis based on live voice or picture is different from a normal doctor-patient contact (10). In some areas telemedicine requires unique techniques. Tele-robotical guaranties differ from what surgeons normally learn. Telemedicine, and distance learning as a prerequisite for it, is lest suited for doctor-to-doctor consultation, and the first contact to a doctor should always be a face-to-face consultation.

Expected outcomes of the project Introduction and Implementation of Distance Learning in medicine are:

- development and integration of informatics-computer technologies in medical education;
- creation of a flexible infrastructure which will enable access to e-learning by all students and teaching staff;
- improvement of digital literacy of academic population;
- ensuring high educational standards to students and teaching staff, and;
- supporting medical staff to develop a "lifelong learning way of life".

The health sector is one of the most evident potential beneficiaries of the Internet "revolution" and World Wide Web resource in the present and in the future, when the tools now available and the system's reliability and efficacy as a whole will be further incremented and improved.

In order to have Distance Learning, a minimum of requirements should be in place: minimal infrastructure: equipment and software, from laptop to notebook, projector and scanner to reliable and quality software, appropriate marketing in Bosnia and Herzegovina environment, education of the educators and administrators, electronic archive (and digital library), support from the officials (moral and financial) and sustainable funding.

Distance learning in medicine is a project started by Prof. Izet Masic, MD, and his team at the Chair of Medical Informatics of the Faculty of Medicine at the University of Sarajevo in the year 2002, first as a pilot study of introduction of distance learning education at biomedical faculties in Bosnia and Herzegovina for improving the educational system in our country (KMIUS). Distance learning is a trend used in many developed countries spreading fast throughout the rest of the world. We hope that this will be just a starting ground for our attempts to improve our educational system. For the first time in B&H, students are able to upload their pictures and other significant data, apply for exams, undertake exams and have their respective results on the website.

Exercises

Task 1

Did you experience e-learning by yourself in the past?

- if yes, write down your experiences (positive and negative);
- if not, write down your expectations of e-learning;
- make groups of three to four students and discuss your experiences or expectations;
- make conclusions of your discussion;
- present these conclusions to other students in a plenary session.

Task 2

Carefully read the theoretical background of this module and the other module on "Public health capacity building: adult education principles and methods" (included also in this book).

- discuss the advantages and disadvantages of tele-education in comparison to other types of education in a group of three to four students,
- make conclusions of your discussion,
- present these conclusions to other students in a plenary session.

Task 3

Using the ICT, try to find out if in your country the distance learning at the university level is available to the students. Present results of your "research" to other students.

References

1. Masic I, Novo A, Kudumovic M, Masic Z. Edukacija iz Medicinske informatike na medicinskim fakultetima u BiH. Acta Inform Med. 2005;13(1):3-8.
2. Masic I, Novo A, Kudumovic M, Masic Z. E-learning at Medical Faculty of University of Sarajevo. Acat Inform Med. 2005;13(3):132-135.

3. Masic Z, Novo A, Masic I, Kudumovic M, Toromanovic S, Rama A, Dzananovic A, Bander I, Masic M, Guso E, Balta E. Distance Learning at Biomedical Faculties in Bosnia and Herzegovina. Connecting Medical Informatics and Bio-Informatics. Proceedings of MIE 2005 (editors: Engelbrecht R, Geissbuhler A, Lovis C, Mihalas G). IOS Press, 2005:267-72.
4. Engelbrecht R, Ingenerf J, Reiner J. Educational standards - terminologies used. Stud Health Technol Inform. 2004;109:95-113.
5. Masic I, Ramic-Catak A, Kudumovic M, Pasic E. Distance learning in the medical education in B&H: E-Health & Education Proceedings, Zagreb, 2002:17.
6. Masic I, Bilalovic N, Karcic S, Kudumovic M, Pasic E. Telemedicine and telematic in B&H in the war and Post war Times. European Journal of Medical Research, 2002; 7(supl.1): 47.
7. Hovenga J. Bricknell L. Current and Future Trends in Teaching and Learning. Stud Health Technol Inform. 2004;109:131-142.
8. Mantas J. Future Trends in Health Informatics – Theoretical and practical. Stud Health Technol Inform. 2004;109:114-127.
9. Mantas J. Comparative educational systems. Stud Health Technol Inform. 2004;109:8-17.
10. Sosa-Iudicissa M, Oliveri N, Gamboa CA, Roberts J. Internet, Telematics and Health. IOS Press, 1997;36:553.
11. UTIC Univerzitetski tele-informatički centar (UTIC) (in Bosnian) [homepage on the Internet]. Available from URL: www.utic.net.ba. Accessed: June 10, 2013.
12. Zimitat C. Designing effective on-line continuing medical education. Med Teach. 2001; 23: 117-122.
13. Harden RM, Hart IR. Centre for Medical Education, University of Dundee, UK. An international virtual medical school (IVIMEDS): the future for medical education? Med Teach. 2002; 24: 261-7.
14. Harden RM. A new vision for distance learning and continuing medical education. J Contin Educ Health Prof. 2005; 25: 43-51.
15. Masic I, Riđanovic Z. Informacijske tehnolgije u medicini i zdravstvu; Acta Inform Med. 1995; 3: 5-8.
16. Google.com [homepage on the Internet]. Available from URL: www.google.com. Accessed: June 10, 2013.
17. Med@Tel. Centar za telemedicinu (in Croatian) [homepage on the Internet]. Available from URL: www.telemedicina.hr. Accessed: June 10, 2009.
18. [homepage on the Internet]. Available from URL: www.medicina.org
19. [homepage on the Internet]. Available from URL: www.publichealth-zagreb.hr.
20. KMIUS Medicinski fakultet Univerziteta u Sarajevu. Medicinska informatika (in Bosnian) [homepage on the Internet]. Available from URL:
http://www.unsamedinfo.org/index.php?NAV=opci_podaci. Accessed: June 10, 2013.
21. http://e-learningguru.com.
22. http://www.dso.iastate.edu/asc/academic/elearner/advantage.html.
23. http://www.medbiq.org/working_groups/virtual_patient/index.html.
24. Web-SP (Web-based Simulation of Patients): http://websp.lime.ki.se/.

Recommended readings

1. Zimitat C. Designing effective on-line continuing medical education. Med Teach. 2001;23:117-122.
2. Harden RM, Hart IR. Centre for Medical Education, University of Dundee, UK. An international virtual medical school (IVIMEDS): the future for medical education? Med Teach. 2002;24:261-267.
3. Harden RM. A new vision for distance learning and continuing medical education. J Contin Educ Health Prof. 2005;25:43-51.
4. Masic I, Ridanovic Z. Informacijske tehnologije u medicini i zdravstvu; Acta Inform Med. 1995;3:5-8.
5. Masic I, Ramic-Catak A, Kudumovic M, Pasic E. Distance learning in the medical education in B&H: E-Health & Education Proceedings, Zagreb, 2002:17.

6. http://www.askinternational.com/knowledge/articles/eBasic/whatsElearn.html.
7. http://www.nsba.org/sbot/toolkit/tiol.html.
8. http://www.babson.edu/News-Events/babson-news/Pages/130107-2012-survey-of-online-learning-results.aspx (Babson Research Study. Retrieved 12 February 2013)
9. Ahmad Zameer, Virtual Education System (Current Myth & Future Reality in Pakistan) (November 16, 2010). Entrepreneurial Tutors. Available at SSRN:
   http://ssrn.com/abstract=1709878
10. The Disadvantages of Online Learning | Elearning Companion. Elearning Companion. N.p.,n.d.Web.12 Feb.2013; http://www.elearning-companion.com/disadvantages-of-online-learning.html
11. Allan B. Blended Learning: Tools for Teaching and Training. Facet Publishing, 2007.
12. Anderson T, Elloumi F. Theory and Practice of Online Learning. Abathasca University, 2004.
13. Chumley-Jones H, Dobbie A, Alford C. Web-based learning: Sound educational method or hype? A review of the educational literature. Academic Medicine. 2002:77S86–S93.S93.
14. Masiello I, Ramberg R, Lonka K. Learning in a web-based system in medical education. Medical Teacher. 2005;27561–563.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Human resource management**
Module: 2.28	**ECTS (suggested): 0.5**
Author(s), degrees, institution(s)	Silvia-Gabriela Scintee, MD, MSc, PhD General Deputy Director, National School of Public Health, Management and Professional Development Bucharest, Romania
Address for Correspondence	Silvia-Gabriela Scintee National School of Public Health, Management and Professional Development 31 Vaselor Str, 021253 Bucharest Romania Tel. +40212527893 E-mail: mailto:gscintee@ispb.ro; sgscintee@snspms.ro
Keywords	Employees training and development, evaluating employee performance, human resource management, human resource planning, job analysis, motivation, recruiting and selection of personnel.
Learning objectives	At the end of this module, students should be able to: • define human resource management and identify its functions; • describe the steps involved in the human resource planning process; • define the term of job analysis; • describe the recruiting and selection processes; • identify methods for employees training and development; • describe methods for evaluating employee performance.
Abstract	This module covers the following topics: Fundamentals of human resource management, Staffing the organization, Employees training and development, Personnel motivation, Employees maintenance.
Teaching methods	Teaching methods include lectures, interactive presentation of key concepts, group discussions, groups assignments, role playing for job interview.
Specific recommendations for teachers	Teachers should discuss the concepts and methods in comparison with the practice in SEE countries. This course takes two hours of lecture and four hours practical session including role-playing. The rest is individual work.
Assessment of Students	1. Reports presented by each group during the working sessions could be considered as part of assessment. 2. A multiple-choice test containing questions from all topics, at the end of the course could be used for final assessment.

HUMAN RESOURCE MANAGEMENT

Silvia Gabriela Scintee

Human resource is the most important asset of an organization, and has a crucial importance for management, as the management is the process of efficiently achieving the strategic objectives of the organization through people. Human resource management is responsible for the people dimension of the organization and is concerned with getting competent people, training them and motivating them to perform at high levels (1). Human resource management is the process that assures the utilization of the employees so that both the organization and the employees to obtain the highest possible benefit.

Some authors make a distinction between human resource management and personnel function. In this view, while human resource management has a strategic role, assuming human resource policies development for the entire organization, the personnel function is supposed to have an operational role, being considered as a tool for human resource policies implementation. Thus, while human resource management is the responsibility of a special department, the personnel management is one of the duties of the managers from all levels in the organization. According to other authors personnel management is the historical term of human resource management and the change appeared with the changing roles of professionals in human resource area (2,3).

Human resource management functions

The main functions of the human resource management are: staffing, training and development, motivation and maintenance.

Staffing function

In order to implement strategies and achieve the stated goals and objectives an organization must be staffed with adequate numbers of properly trained personnel. The staffing function is a continuation of strategic planning process, when after determining how goals and objectives will be attained the managers should determine what jobs need to be done and by whom. These activities are included in the human resource planning process that determines staffing needs. The outcome of this process will be either recruitment or decruitment (1). When acquisition of personnel is needed, the manager should gather two types of information: information on the job and information on the persons eligible for the job, than to match skills, knowledge and abilities to required job. When manager has to sever people from the organization, he/she should engage in activities to assist and monitor exit.

The staffing function includes: human resource planning, job analysis, job description, recruitment, selection, integration of the new employee, assistance and monitoring personnel exit.

Human resource planning

Human resource planning is the process by which an organization ensures that it has the right number and kinds of people able of efficiently and effectively performing the tasks required for achieving the strategic objectives of the organization (1,4). As organizations are dynamic, permanently under the influence of external environment and internal factors, planning should be a continuous process. The main steps of human resource planning are: strategic objectives analysis, estimating staffing needs, assessment of current human resources, forecasting changes in the present workforce and development of an action plan.

Strategic objectives analysis. Strategic objectives are broad statements that establish targets the organization will achieve in a certain period of time. The analysis of the current strategy determine how goals and objectives will be attained and to what extent the organization can meet its objectives given the internal strengths and weaknesses and external opportunities and threats. Commonly referred to as SWOT analysis, this managerial tool could bring information on what skills, knowledge and abilities are available internally, and where the shortages in terms of people skills or equipment may exist (1). The analysis should take into account influencing factors such as: anticipated demand for services, changes in professional practice or labour supply, development of new technologies.

Estimating staffing needs. The analysis of the objectives and of the ways in which they will be attained gives information on the number and types of the jobs needed and the skills and knowledge required for the jobs. Unless drastic changes will occur, such as reengineering initiatives, major organizational changes, human resources needs estimates could be made for a certain time period, using the established staffing ratios for most major functions. For example, for the further development of an outpatient department at a hospital, projection of the needed staff can be made (2). Using the projected volume of service and the accepted nurse staffing ratio, it can be determined the number of nurses needed.

Assessment of current human resources. Assessment of current human resources is a valuable input in the human resource planning process by determining what skills are currently available in the organization and how are they used. Each organization should generate a detailed human resources inventory report listing all employees by title of the job, education, training, prior employment, performance ratings, salary level, languages spoken, abilities and specialized skills (2). Such an inventory could be also used for other activities such us internal recruitment or selecting individuals for training and development.

Forecasting changes in the present workforce. An organization will experience turnover through retirement, death, voluntary separation, etc. Based on historical data the changes in the present workforce could be forecast. For managerial positions some organizations use the replacement chart (1). This is a diagram that determines if there is within organization a sufficient managerial potential to cover future vacancies. The main information listed for each individual is: the current position, expected replacement time, and possible replacements with their potential and readiness in occupying the job.

Development of an action plan. The previous steps bring the necessary information for developing an action plan to fill the staffing needs through recruiting and hiring, transferring or enhancing the skills of existing employees by training and development. The action plan will specify: the jobs to be created, transformed or cut off, the implications at institutional level, the number of persons to be hired and specifications of their characteristics, the movement of the personnel within the organization and the training needs, the methods of sorting out the unpredictable losses, the costs of covering staffing needs and the timetable of each activity.

Job analysis and job description

Job analysis is a systematic examination of the activities within a job (1). This analysis involves the description of the job content (the goal of the activity, tasks to be fulfilled, duties and responsibilities, resources used, expected results), what are the job requirements (knowledge, abilities, skills required), what are the working conditions (physical environment, hazards), and the social environment (individual or group work, communication skills required, relationships to other jobs). There are three basic methods for job analysis: observation questionnaire and interview. Observation provides firsthand information. This can be done directly or reviewing films of workers on the job. Observation can not bring exact information as people being watched act differently than in their day to day activity. The interview has an increased accuracy in assessing jobs by involving employees in analysis. In order to increase the effectiveness of this method it is recommended a combination between individual and group interview. Structured questionnaires could also be used for gathering information about a job. The disadvantage of this method is that exceptions to a job may be left out and there is no opportunity to ask follow-up questions. Other methods that can be used for job analysis are technical conference (specific job characteristics are obtained from experts) or diary method (workers are asked to record their daily activities). The main purpose for job analysis is to gather information in order to develop: job descriptions, job specifications and job evaluation (1).

A **job description** is a written statement of what the jobholder does, why and under what conditions. The content and format of job descriptions vary among organizations. Yet, the general job description format include (1,5):

Job description format

Name of organization
Name of division/department
Job title
Grade of job
Job purpose
Duties to be performed
Authority and responsibilities of the jobholder
Supervision given or received
Relationships with other jobs
Environmental working conditions
Special provision (e.g. confidentiality)
Terms and conditions (e.g. salary, working hours, holidays)

Job specifications states the characteristics that the jobholder must posses in order to perform the job successfully. These characteristics are identified also during job analysis and refer to the knowledge, skills, education, experience, certification and abilities needed to do the job effectively. Job description and job specifications are used in activities such as human resource planning, recruitment, selection, performance evaluation, compensation plans.

Job evaluation is the process of determining the value of each job in relation to the other jobs within organization. On the basis of job evaluation the jobs in an organization are ranked and placed in a

hierarchical order (3,6). The resulted ranking should be used in order to establish the compensation programme.

Recruitment and selection

Recruitment is the process of searching and attracting potential candidates for present or anticipated vacancies. The recruitment sources could be either internal or external (2). The internal search attempts to identify present employees who can fill a vacancy by transfer or promotion. This method is cost effective, quick and motivating for the employees. The external search is done mainly by advertisements that can be placed in different newspapers, magazines, electronic sites or public places, depending on the type of the job. The main elements to be included in a vacancy announcement are given below (6):

Vacancy announcement elements:

a. Organization name
b. Title of the job
c. Location of the job
d. Employment duration
e. Description of duties
f. Job specifications
g. Salary and employment terms
h. Application procedure

Other external sources are employment agencies, schools, colleges, universities, professional organizations or even unsolicited applications. The selection of the recruitment source depends on the job characteristics, labour market supply, and geographic workforce distribution. The success of the recruitment process is influenced by factors such as: organization reputation, the attractiveness and nature of the job, internal policies of the organization, legal requirements, and the recruitment budget (4).

Selection. The next step in acquisition of personnel is to choose from all qualified applicants for a job identified through recruiting the "right" one. This is a very important decision, as a good selection process can save costs for personnel replacement or training and can increase the work productivity. There could be considered two steps of the selection process: initial screening and final selection. Initial screening consists in gathering preliminary information about candidates and excluding those who are not suited for the job in terms of training, experience and ability. Among methods used for initial screening are curriculum vitae, intention letter, application form, letter of recommendation, employment tests.

Curriculum vitae (CV) brings information mainly on training and experience, but also give some insights about candidate personality, when looking at its clarity, stile and logic sequence of ideas (6). The information to be included in a CV is:

- personal data and characteristics (name, surname, contact details, date of birth, nationality, marital status);
- education (institutions, dates, degree or diplomas obtained);
- present position (company, location, description of main tasks);
- work experience (employment record, institutions, dates, main tasks and responsibilities);
- scientific activity (papers, presentations, publications);
- other skills (e.g. proficiency in foreign languages, computer literacy);
- other information (if appropriate, e.g. hobbies, preferences for leisure time).

Intention letter. While CV is just an inventory of the person's history regarding training and experience, the intention letter is the mean by which the candidate exposes his motivation and desire to get the job. The intention letter also talks about candidate's professional and human qualities and about his compatibility with the job. No longer than one page, an intention letter should not contain the information from CV, but has to wake up the reader interest in setting an appointment for the candidate.

Application form. There is not a general format for the application forms. Each organization has its own format, some of them requiring may be only applicant name, address and telephone number, others requesting the completion of a more comprehensive profile. In general, application forms bring less information than a CV, but they are very common, representing a standardized tool for information gathering which makes comparison between candidates easier (6). Some application forms could include statements giving the employer the right to obtain previous work history of the candidate, to dismiss him for falsifying information or to end the work relationship at the employer will. If the candidate does not sign such a form his application is removed from consideration (1).

Letter of recommendation. Information about candidates could be obtained from other persons, too. Even criticized as being subjective, recommendation letters are still very common. They depend on the intention and the degree of information of the person who issues it. Usually there are requested two or more recommendation letters.

The initial screening shortlists the candidates for the final selection. Both shortlisted and not shortlisted candidates should be announced about their results through an official letter. The shortlisted candidates are asked to come for the final selection that can be done by employment tests or by interview.

Employment tests. Tests are used mainly for two purposes: the assessment of the candidates' knowledge, abilities and skills, and the psychological evaluation of the candidates. The second category is given more importance as many studies have shown that the employees' performance is more related to their personality characteristics than to the knowledge they have. There are hundreds of tests that can be used by an organization in selection purposes. They are measuring intellect, memory, perception skills, spatial ability, motor ability, personality traits, etc. bringing information that can not be obtained from the candidate and that can make predictions on the person behaviour. Tests can be written tests or simulation tests. The last ones require the applicant to engage in specific activities and behaviours necessary for doing the job.

Assessment centres. An organization can also address for initial screening of its candidates to an assessment centre. These are specialized institutions that combine more methods in selecting candidates. All applicants are received at such a Centre for a 2-4 days period, being subjected to individual and group testing by: interviews, solving problem exercises, group discussions, role playing, personality and general ability tests, etc. In the same time it is assessed the candidates social behaviour (1).

Interview. Interview is almost universal accepted as the final selection tool, evaluating the candidates' compatibility with the job, motivation and abilities of integrating themselves in the organization. The interview gives the opportunity of clarifying the previous gathered information on the candidate and also can test the candidate reaction under particular situations such as stress, conflicts, etc. The interview's validity and reliability are subject of criticism (1). In order to increase the effectiveness of the interview, it should be conducted by a person familiar with the interview technique and having some specific qualities: determination, discipline, self-control, tolerance, empathy, lack of prejudices. An interview should be careful prepared, paying attention to the place where it will be held, obtaining detailed information about the job and its requirements, studying applicants information gathered in the initial screening stage, planning time, developing guidelines for interview and a list of criteria to be evaluated during the interview (5,6).

Integration of the new employees

After selection the new employee is helped to integrate in the organization, in order to become productive as soon as possible. The human resource department is responsible with enrolling new employee in benefit plans, issuing an identification badge/card. The chief of the department in which the new employee will work will take care of preparing the work place, and will delegate a supervisor to prepare and implement an orientation programme (4). The supervisor will introduce the new employee to other colleagues, will explain the organizational structure and function, will explain in detail the department specific work methods and internal norms and rules. The supervisor also helps the new employee to gain acceptance by others and will morally support him with any personal problems. Usually, in a month time the manager will meet the new employee in order to evaluate the extent to which he integrated in the organization.

Assistance and monitoring of personnel exit

Sometimes the employees have to leave the organization from various reasons. The personnel exit should also be assisted and monitored. Besides activities like completing personnel records, collecting employer-provided equipment and processing final pay, a manager could involve in activities oriented to the alleviation of psychological impact of leaving the job and to assisting employees in finding employment (2). Thus, some organizations have a preretirement programme consisting in preparing employee for the psychological, emotional and financial changes in retirement. When jobs are eliminated for various reasons (changing demand, downsizing, mergers, etc) the leaving employee should receive an earlier notice and should be helped in finding a new working place. Also, the employee could be tested for discovering abilities for other jobs and helped in the process of professional re-orientation and re-location.

Training and development function

Training and development is a key element in helping employees to maximize their potential. The goal of training and development function is to have competent employees who possess the up-to-date skills, knowledge and abilities needed to perform their current jobs more successfully. Although there are similarities in the methods used to affect learning, the terms of education, training and development are different (6). Education refers to a basic teaching, a long term learning process, directed to obtaining knowledge, abilities and skills that allow individuals or groups to perform the social roles. Education is focused mainly on individual needs and also on community needs. Training is a learning process oriented to the acquiring of specific knowledge, abilities and skills necessary to the individuals or groups for performing a job. Training is job or tasks oriented, it has a continuous character and it might assume

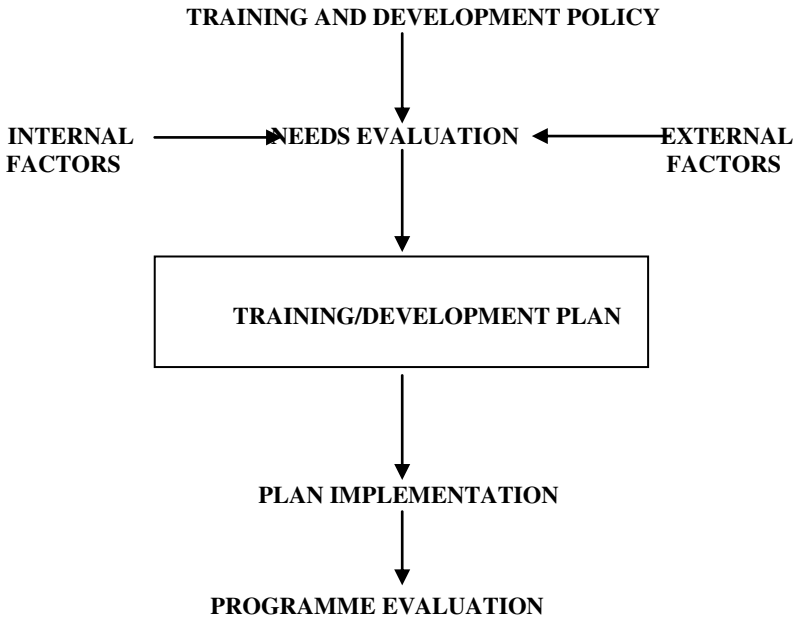
changing of skills, attitudes or behaviour in order to immediately adapt to the present job requirements. Development is a learning activity oriented rather to the future needs than to the present ones. Employee development focuses on the future jobs in the organisation and career progress for which new skills and abilities will be required.

Each organisation should have a continuous training and development programme. Specific training and development needs are given by:

- hiring new employees;
- acquisition of new technology and equipment;
- low performance of the organisation;
- occurrence of some events with a higher frequency than usual (e.g. nosocomial infection in a hospital);
- changing demand for services;
- organisational changes.

The development of a training and development programme has the following steps (see figure 1): policy development, needs evaluation, training/development plan, plan implementation and programme evaluation (6).

Fig. 1 Steps of a training and development programme



Training and development policies are included in overall human resource policies of an organization and have to be in accordance with its general policies (6). Training and development programmes should take into account the training and development policies that usually state the organization's commitment of assuring to the employees the appropriate means for training in order to successfully perform their jobs.

Elaboration of a training and development programme should be preceded by training and development needs assessment. The training need is represented by the deficit of knowledge, abilities and skills in relation to the level required by the job or by the organizational changes (3,6).

The main information sources for needs assessment are:

- the organization – we will look at the organization's goals, structure and functioning;
- the job – what tasks have to be completed to achieve the organization's goals, what are the requirements for effectively performing the job;
- the employee – what is the level of employee's performance, what are the deficiencies he has in the skills, knowledge or abilities required to perform the job;

Training and development needs assessment has to take also in account all internal and external factors that might contribute to the changing of the organization needs.

Once it has been determined that further training and development is necessary, an action plan will be developed (6).

The structure of the training/development plans:

Training goals
Training objectives
Target groups
Training content

Training methods
Time schedule
Estimated necessary resources
Evaluation and monitoring tools

Training goals should be clearly stated and they can refer to: increasing capacity for problem solving, enhancing ability for performing specific activities, acquiring skills for performing new tasks, increasing communication skills, modifying attitudes towards change. Training objectives should be tangible, verifiable, timely and measurable. They have to reflect the real changes in the employees' knowledge, abilities, skills or attitudes. Training content will be established in accordance with the training objectives and the level of previous training of the target group.

Training methods can be classified as either on-the-job or off-the-job training (1). On the job training method is the most used, being simple and less expensive. It is a learning by doing method, placing the employee in actual work situations and asking him to do the tasks. This method is more appropriate for jobs that are difficult to simulate or for those that can be easily learned by watching and doing. Examples of on-the-job training are:

- apprenticeship – is used for training in different trades where skills are so complex that can not be acquired on theoretical basis or by simulation. It consists in putting the trainee under the guidance of a skilled master;
- job instruction training – consists in explaining the trainees what they are suppose to do, verifying their understanding and placing them in the job under a supervisor to call upon if they need assistance.

On the job training has the risk of low productivity, but has the advantage of motivating workers, increasing employee morale and understanding.

Off-the-job training has a various number of techniques:

- lectures – designed to communicate theoretical concepts, to describe tools or to present technical, problem-solving skills;
- seminars and workshops – for more interactive discussions and practical exercises in which to apply theoretical knowledge;
- simulation exercises – in which trainees are performing different tasks in a working like situation; this also may include: case studies, role playing, group decision-making, computer based simulation, training on real equipment away from the work setting;
- videos and films – use media production to demonstrate specialised skills that can not be easily presented by other methods.

Developing methods can also take place on-the job or off-the-job (1). Among on-the-job techniques there are:

- job rotation – consists in moving employees to various jobs in the organisation, either on horizontal or vertical, with the purpose of expanding their skills, knowledge and abilities. This method gives the employee an overall view on the organisation activities, turns him from a specialist to a generalist, avoids boredom and stimulates the development of new ideas;
- working as staff assistant – the employee works as the "shadow" of an experienced person from the next higher level. Working as an assistant, the employee has the opportunity to be exposed to the whole range of the activities in that position, he learns by performing many duties under direct supervision and get used with assuming the duties and responsibility of the higher level;
- committee assignment – the employee is appointed to temporary or permanent committees. This allows employee to take notice about specific organisational problems and to learn from the others example how to solve different problems and to participate in decision making.

Off-the-job methods could be done by traditional forms of instruction such as lectures, seminars, simulation exercises or by modern techniques like outdoor training.

- lectures and seminars – they are offered for acquiring knowledge or for developing employees conceptual and analytical abilities and could be organised either in class or by distance learning;
- simulations – as seen above, simulations are exercises in which employees are performing different tasks in a working like situation. The most used are: case studies, role playing, decision games;
- outdoor training – also called wilderness or survival training, this method teaches the importance of working together and involve emotional and physical challenge. The most known techniques are: white-water rafting, mountain climbing, paint-balls games or surviving one week in the jungle.

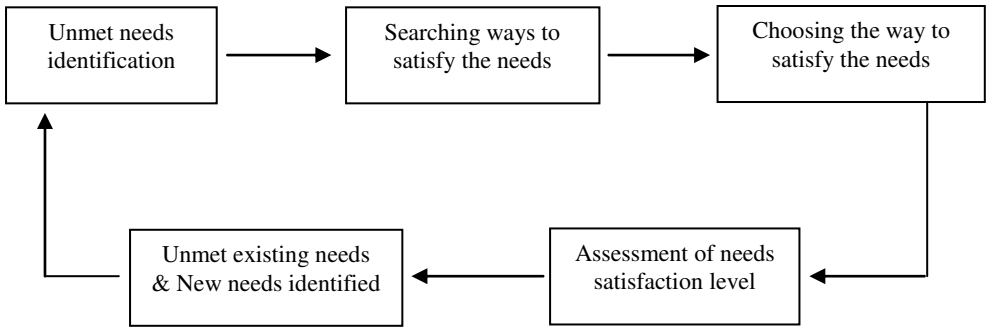
Motivation function

Motivation is a key determinant of employees' performance. The concept of motivation is based on the way in which people are given attention and on the feelings that they have in relation with their work. To motivate employees means to satisfy their unmet needs, to stimulate them to work better in order to achieve the organization's goals. Unmet needs cause discontent which is reflected in employee's negative

behaviour and attitudes, producing tension and low productivity. The motivation process is cyclical (2). It starts from identifying individuals unmet needs, after that ways to satisfy the needs are searched for and the most convenient is chosen. The needs satisfaction is recommended to be followed by the assessment of needs satisfaction level, which may confirm the failure of satisfying the need or identifies a new need and the cycle is restarted (figure 2).

The motivation theories. The multitude and diversity of theories developed to explain human motivation reflects its complexity. The most important motivation theories can be divided in two categories: content theories and process theories (2,3,5). While content theories focus on "what" motivates people, process theories focus on "how" motivation is initiated and sustained.

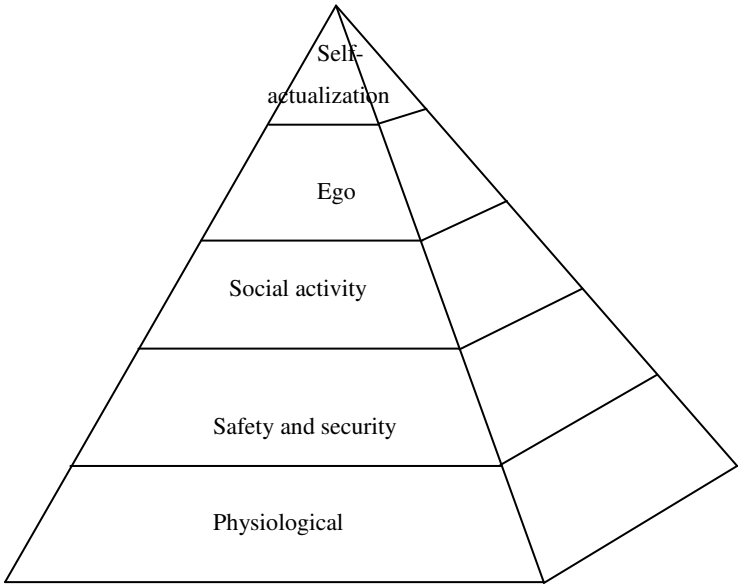
Fig. 2 The motivation process cycle



Among content theories there are:

Maslow's Hierarchy of Needs– considers people needs arranged in the following hierarchy (from lowest to highest): physiological, safety and security, social activity, ego and self-actualization (figure 3).

Figure 3. Hierarchy of people needs according to Maslow Theory



According to this theory, only needs not yet satisfied influence behaviour and once the needs from a level are fulfilled, the individual moves up to the next level. Primarily, an individual has physiological needs, such as air, water, food, shelter and sex are basic for an individual. Once these survival needs are met, the individual turns to the next level: safety and security, represented by needs for health insurance and other benefits that ensure protection against physical harm and deprivation. The third level of needs includes the need for belonging, friendship affection and love. Examples of ego needs are: the need for independence, achievement, recognition, self-esteem and status. In the top of the hierarchy are self-actualisation needs represented by continuing growth and development, opportunities for self-expression and self-fulfilment.

Alderfer's ERG Theory – refers to the three categories of needs: **E**xistence needs – including material and physical needs that can be satisfied by air, water, money and working conditions, **R**elatedness

needs – that involve other people and are satisfied by social and interpersonal relationship and **G**rowth needs – including all needs satisfied by an individual through creative or productive contributions. Similar to Maslow's theory, people focus first on needs that are satisfied by more concrete ways.

***Herzberg's Two*-Factor Theory –** says that job satisfaction consists of two separate and independent factors: intrinsic job factors such as responsibility and recognition which motivate when they are adequate and extrinsic factors called also "hygiene factors" that only placate employees when are present but they cause dissatisfaction when they are deficient. Among the hygiene factors identified by Herzberg there are: organisational policy and administration, interpersonal relationships, salary, job security, working conditions.

***McClelland's Learned Needs Theory* –** states that people learn about their needs through life experiences and there are three major needs in workplace situations: the need for achievement, the need for power and the need for affiliation.

The most known process theories

***Vroom's Expectancy Theory* –** based on the concept that people have preferences for outcomes and if they have a strong preference for a particular outcome they will attach to that outcome a high valence.

***Adam's Equity Theory* –** focus on people desire of being treated fairly and states that individuals assess whether rewards are equitably distributed within organisation by calculating the ratios of their efforts to the rewards they receive and compare them to the ratios for others in similar situations.

***Locke's Goal Setting Theory* –** affirms the importance of goals in motivation as people focus their attention on the concrete tasks that are related to attaining their goals and persist in the task until the goals are achieved.

All the above mentioned theories are based on the McGregor observation of the importance of managers' attitudes about people in determining their approach to motivation. In 1960s Douglas McGregor proposed two alternative sets of assumptions that managers hold about human nature in workplaces: Theory X – according to which managers view people in negative ways and Theory Y – that argues that managers view people in positive ways. According to McGregor Theory Y assumptions are more valid than Theory X and employee motivation would be maximised by giving workers greater job involvement and autonomy.

Motivation principles (1)**:**

- Put the right person in the right place. No reward or stimulating factor could increase a person's productivity if that person lacks the ability to perform the job. Matching properly the employee to the job should be an objective of recruiting and selection.
- Managing by objectives. People work better when their activity is goal-directed because this is challenging and it is clear to the employee what is to be done. The results are even better if individual objectives are mutually set and are in accordance with the department and organisation objectives. Continuous feedback is also important for increasing individuals' performance.
- Understanding individual needs. Individuals are different, and each individual has its own set of needs. So, unmet needs assessment should be done for each employee.
- Individualise rewards. As the individual needs are different from a person to another, rewards should also be different. What motivate an individual could not be motivating for another one.
- Reward performance. The best way to encourage increasing performance is to reward individuals for their performance or to relate any other reward they receive with the achievement of the organisation goals.
- Use an equitable rewarding system. People are concerned not only with the rewards they receive, but also with the equity of their rewards compared to what others receive. So efforts must be made in order to ensure that the reward system is fair, consistent and objective.
- Money is the best reward. As money is the main reason for people to work, no other reward would be appreciated if they were not paid sufficiently to cover their basic needs.

Possible ways to increase motivation

- Job related rewards
- Job characteristics such as: the requirement for using various skills and talents, the requirement of completion of a whole and identifiable piece of work, the impact the job has on the lives or work of other people, a high degree of autonomy, a high degree of information received back on the effectiveness of his/her performance.
- Job enrichment – the worker is allowed to assume some of the tasks executed by his/her direct manager.
- Job rotation – the employee has the opportunity to diversify its activity and offset the occurrence of boredom.

- Work at home – this affords employee, especially women, to combine both their careers and family responsibility.
- Flexible hours – increases workers' freedom; employees assume responsibility of completing a specific work in a specific time, and this increases their feeling of self-worth.
- Training courses.
- Assuring a safe, pleasant and practical working environment.
- Rewards not related to the job
- Tangible rewards: premiums, stocks, insurances, presents, free lunches, free snacks and coffee at the break, etc.
- Social rewards: free tickets for spectacles, picnics, trips, free access to company clubs, etc.
- Acknowledgements – diplomas, certificates, mentionings in the panel of honour, informal acknowledgements.

Maintenance/retaining function

Another function of human resource management is to put into place activities that will help retain productive employees. These activities includes: appraising employees performance, moving employees within the organization through promotion or transfer, providing employee assistance and career counseling, administering compensation and benefits, ensuring a healthful workplace and personal safety (2).

Performance appraisal evaluates an employee's work by comparing actual with expected results. It should be done at any level, from employees and managers.

Uses of performance appraisal (6)

- to collect information in order to evaluate if work results are those expected and, if not, to determine why not,
- to help decision-making in regards with compensation schemes and other benefits,
- to determine the further use of the employee (if he/she should stay at the same work place, or should be transferred, promoted, demoted or deployed),
- to evaluate training needs by identifying areas in which performance could be increased in proper training is undertaken,
- to motivate employees for working better by providing feed-back and making the results available to the others, too,
- to increase communication between employee and supervisor, allowing the opportunity to discuss the problems that are responsible for a low performance, and
- to provide information on employee assistance and counseling needs.

Performance appraisal principles

- Evaluation criteria should be formulated according to job description. They have to be clearly stated, easy to measure and in small number. Examples of evaluation criteria are: the degree of fulfilling with tasks, the degree of assuming responsibility, initiative and creativity, etc.
- The measuring of performance should be done against specific standards. These are established by job analysis, which gives information on the tasks that have to be fulfilled, the way in which the tasks should be performed. The performance standards cover: the quantity and quality of work, the efficient use of resources in order to maintain costs, the compliance with the time schedule, the specific requirements for the job (such as team work abilities, flexibility, communication skills, etc).
- The appraisal should be prepared and scheduled in advance. The employee should have permanently access to his job description, which should also have attached the list of performance appraisal standards and the schedule of the periodical evaluations. Thus the employee has the opportunity to prepare in advance. On the other hand, the manager should be prepared in advance reviewing the employee's job description and his previous performance measures.
- The employee should be involved in appraisal by taking active part in the discussion, raising questions, adding his own perceptions about his work and also by a self-evaluation.
- The employee should be familiar with the purpose of the appraisal and the evaluator should behave such a way that the employee understand that the appraisal has the role of helping him and not of punishing him.

Performance appraisal methods can be classified according with the approach in: methods based on absolute standards, methods based on relative standards and methods based on objectives (1,7).

Among the methods based on absolute standards there are:

- Essay method – the appraiser writes a narrative description of employee's strengths, weaknesses, potential and suggestions for improvement. This method can provide specific information, but makes difficult the comparisons between individuals.

- Critical incident method – looks mainly at behaviours, focusing on those critical aspects that make a difference between doing a job effectively and doing it ineffectively. The comparison and ranking of employee is difficult by this method.
- Checklist – the evaluator uses a list of behavioural descriptions and checks off those behaviours that apply to the employee. The list is evaluated by another person and this reduces some bias as the rater and the scorer are different persons.
- Rating scale – the most common method, it can be used to assess job dimension attributes such as quantity and quality of work, job knowledge, or personal traits and behaviours such as cooperation, dependability, loyalty, attendance, honesty, attitudes, initiative. For each scale there is a scoring mechanism using descriptive adjectives from "poor" to "excellent" or numerical values that often range from 1 (poor) to 10 (excellent).
- Forced-choice method – is a special type of checklist where the rater must choose between two or more statements, each statement being favorable or unfavorable. The appraiser will identify which statement is most/least descriptive for the individual being evaluated.

In the category of appraisal methods based on relative standards there are:

- Group order ranking – this requires the evaluator to place employees into a particular classification, such as "top 10". This method prevents raters from inflating their evaluation by rating everyone as good.
- Individual ranking – requires the evaluator to list the employees in order from highest to lowest.
- Paired comparison – it ranks each individual in relationship to all others on a one-on-one basis. Each person is scored by counting the number of pairs, among his colleagues, in which he is preferred member.

The third approach to performance appraisal makes use of objectives, being commonly referred to as management by objectives (MBO). It consists in four steps: goal setting, action planning, self-control and periodic reviews. For each employee specific objectives are established jointly by the supervisor and the employee, and also realistic plans are developed in order to attain the objectives. The employee is monitoring and measures his/her own performance, with periodic progress reviews done by supervisor.

Performance appraisal errors. The main problem with the performance appraisal methods is that all of them allow some bias (1,7). The most common errors are described below:

- Leniency/severity errors – the individuals within an organization are evaluated by different persons. Some evaluators are more generous than others, so the performance is evaluated either higher or lower than it really is and comparisons between individuals are not reliable.
- Halo effect – the evaluator's general opinion on an employee is influenced by a single specific aspect.
- Central tendency – is the evaluator tendency of avoiding the extremes and rating everyone in the middle.
- Similarity error – is given by the fact that the evaluator rates other people in the same way that they perceive themselves, by projecting those perceptions onto others.
- Other errors are given by: prejudices, different cultures, recent influencing events.

In order to reduce appraisal errors a combination of two or more methods is recommended.

Exercises

Students will perform four exercises, after each introductory lecture, divided into small groups. Total time required for exercises is four hours.

Task 1: Small group discussions and reports (45 minutes)
Recommended subjects for group discussion:
Human resource planning, advantages and limits.
The possible recruitment sources for managerial jobs in health sector.
How doctors in your country are best motivated.
How performance is assessed in different organizations in your country (from students experience or after visiting some organizations and collecting information)

Task 2: Developing skills in human resources management (45 minutes)
Recommended assignments for group work:
Prepare the following documents and report:

- Job description
- Write the job advertisement
- CVs and intention letters
- Find different blank application forms from different organizations and compare them. Which application form is most preferred overall by the class.

- Develop a training plan for middle level hospital managers

Task 3: Web-wise exercises (45 minutes)
Search the web to identify current job opportunities and report:
http://www.careermosaic.com
http://www.occ.com
http://www.who.ch
Look also for other sites.

Task 4: Role play (65 minutes)
Job interview:
Choose up to seven applicants for a certain job who will submit their CVs and intention letters for applying to a job. A small group (4-5 persons – the interview commission) will shortlist 2-3 candidates for interview. Than the interview will be conducted and it will be chosen the best person for the job. The other persons in the class will discuss at the end the positive and negative aspects observed during the interview.

References

1. De Cenzo A. David, Robbins P. Stephen, Human Resource Management. 6th ed. New York: John Wiley & Sons Inc. 1999.
2. Longest JBB, Rakich JS, Darr K. Managing Health Service Organizations and Systems, 4th ed. Baltimore: Health Professions Press 2000. pp. 531-586, 771-800.
3. Armstrong M. A Handbook of Personnel Management Practice. London: Kogan Page 1996.
4. Higgins M J. The Management Challenge. 2nd ed.. New York: Macmillan Publishing Company 1994. pp. 419-458.
5. Mullins JL. Management and Organisational Behaviour. 4th ed. London: Pittman Publishing 1996. pp. 625 –706.
6. Cole GA. Personnel management, 3rd ed. London: DP Publications Ltd 1993.
7. Pitaru HD. Managementul Resurselor Umane. Masurarea Performantelor Profesionale, Bucuresti Ed. ALL, 1994.

Recommended readings

1. Leung WC. Managers and professionals: competing ideologies. BMJ 2000;321:2.
2. Singh D. Hospitals must use London's appeal to tackle workforce shortages. BMJ 2003;327:70.
3. MacDonald R. Implementing the European working time directive. BMJ 2003;327:9-11.
4. Dosani S, Adsett J. The staff grade dilemma. BMJ 2002;325:170 s.
5. West M. How can good performance among doctors be maintained? BMJ 2002;325:669-670.
6. Edwards N, Kornacki MJ, Silversin J. Unhappy doctors: what are the causes and what can be done? BMJ 2002;324:835-838.
7. Leung WC. Studying for an MBA. BMJ 2000;320:2.
8. Cockcroft A, Williams S. Staff in the NHS. BMJ 1998;316:381.
9. Delva MD, Kirby JR, Knapper CK, Birtwhistle RV. Postal survey of approaches to learning among Ontario physicians: implications for continuing medical education. BMJ 2002;325:1218.
10. Stockdale S. How to stay motivated with an impossible workload. BMJ 2002;324:141.
11. King J. Giving feedback. BMJ 1999;318:2.
12. Ham C. Improving NHS performance: human behaviour and health policy. BMJ 1999;319:1490-1492.
13. King J. Dealing with difficult doctors. BMJ 2002;325:43.
14. Wood LEP, O'Donnell E. Assessment of competence and performance at interview. BMJ 2000;320:2.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Health Databases, Health Indicators and Health Reporting**
Module: 2.29	ECTS 2.5
Author(s), degrees, institution(s)	Doris Bardehle, Prof. Dr. sc. med., Berlin School of Public Health, Charite, Coordinator of the Scientific Board of the Foundation of Men´s Health, Berlin, Germany
Address for Correspondence	Max-Koska-Straße 7 13189 Berlin, Germany
Keywords	Health Data Bases, Health Indicators, Health Reporting, Gender-specific indicators.
Learning objectives	Overview about the objectives and construction of databases, use of classifications for a health indicator set, use of indicators for health reporting and monitoring of health targets.
Abstract	Health Databases and Health Indicator Sets have been developed from 1980 and onwards. The first global database is the Health for All Database of WHO, European Office for Europe, developed after the Alma Ata conference 1978. The Global International Databases are used for the targets of WHO, of United Nations (Millennium Development Goals) and European Union (EUROSTAT databases and ECHI Health Indicator Set of the EU, DG V) and OECD-database for international trends in Economy, and in Health Care Systems, too. The South Eastern European Countries developed a Minimum Health Indicator Set (MHIS) and demonstrated, that such Regional Health Indicator Set works well for comparison of development in the health status and health care system, but if there is no promoter to continue the work with the health indicator set, a break-down is the consequence. In the second part of the lecture cycle there will be explained the scientific background of health indicator sets, classification and evaluation methods as well as the categories of indicator sets and the data sources where the indicators are taken from.
Teaching methods	Two-three Lectures, Search in Databases, Skill-training – Build-up a key indicator set for a country/region, presentation of a topic relevant for health reporting.
Specific recommendations for teachers	Give an overview about the development of international and national databases and the growing use of health indicators and graphic tools for health reporting. Composite indicators like Burden of disease calculations, DALY as Golden Standard for national health indicator sets should be explained and trained during seminars and exercises. Explain the inclusion and exclusion criteria for indicators and the need for quality check and metadata base description of each indicator. Explain the students the different data sources and the advantage and disadvantage of different data sources.
Assessment of Students	

HEALTH DATABASES, HEALTH INDICATORS AND HEALTH REPORTING

Doris Bardehle

Background

Most important sources for health databases and health indicators

This module follows the lines of the documents of:

- United Nation with Department of Population Statistics and Millennium Development Goals Program;
- Organization of Economic Co-operation and Development (OECD);
- World Health Organization (WHO), Headquarter (Geneva);
- World Health Organization (WHO), Regional Office for Europe (Copenhagen);
- European Commission and European Union, as well as EUROSTAT.

The European Union Health Databases and Health Indicator sets base upon:

- The former Health Monitoring Program (HMP) (1998-2003) which created the "Set of Community Health Indicators (ECHI)";
- The New Public Health Program (2003-2008) (19);
- The European Union Health Strategy 2008-2013 (20), the Second Program of Community Action in the Field of Public Health;
- The Health Monitoring System (2009-2013) with the creation of HEIDI Data tool for storage of ECHI Indicators (26), as well as;
- The Europe 2020 Strategy: Investing in Health (24).

Health Policy and types of different health databases and indicator sets

United Nations, Department of Economic and Social Affairs, Statistics Division, Millennium Development Goals (MDG)

At the Millennium Summit in September 2000, the largest gathering of world leaders in history adopted the UN Millennium Declaration, committing their nations to a new global partnership to reduce extreme poverty and setting out a series of time-bound targets, with a deadline of 2015, that have become known as the Millennium Development Goals (MDGs) (1,2) (http://www.unmillenniumproject.org, downloaded: 05.03.2013).

The eight goals for 2015 are the following:

1. Eradicate extreme poverty and hunger
2. Achieve universal primary education
3. Promote gender equality and empower women
4. Reduce child mortality
5. Improve maternal health
6. Combat HIV/AIDS, malaria and other diseases
7. Ensure environmental sustainability
8. Develop a global partnership for development.

The MDGs are a positive step toward global health. The latest progress report was published 2012 in "The Millennium Development Goals Report 2012" (2,3) (http://mdgs.un.org/unsd.mdg). This report highlights some milestones. Some of the 8 goals, 21 targets and 60 indicators have been reached, i.e. the reduction of extreme poverty by half or the better access to drinking water. The infant and maternal mortality have been decreased, but slower than expected. For the years 2010-2015 the most important goal is the eradication of extreme poverty and hunger (<1$ per day income per person). The Public Health goals are "Reduce child mortality", "Improve maternal health" and "Combat HIV / AIDS, malaria and other diseases" with 22 health indicators. You will find these indicators in the MDG Reports and in the WHOSIS: The WHO Statistical Information System (4) (www.who.int/whosis/data).

Organization for Economic Co-operation and Development – OECD

The OECD comprises 34 member states, of which 25 European countries and Canada, Mexico, United States, Australia, Japan, Korea, New Zealand, Chile and Israel (www.oecd.org/documents/explanatory_notes; www.oecd.org/statistics).

The conditions for membership are a high level of economic regulations and high level of statistics, incl. health statistics. OECD uses its wealth of information on a broad range of topics to help the governments foster prosperity and fight poverty through economic growth and financial stability. OECD

helps ensure the environmental implications of economic and social development taken into account. OECD Factbook 2013 is a comprehensive and dynamic statistical annual publication from the OECD. More than 100 indicators cover a wide range of areas: agriculture, economic production, education, energy, environment, foreign aid, health, industry, information and communications, international trade, labor force, population, taxation, public expenditure, and R&D. In 2013, the OECD Factbook features a focus chapter on gender.

Data are provided for all OECD member countries including area totals, and in some cases for selected non-member economies (including Brazil, China, India, Indonesia, Russia & South Africa). For each indicator, there is a two-page spread: a text page includes a short introduction followed by a detailed definition of the indicator, comments on comparability of the data, an assessment of long-term trends related to the indicator and a list of references for further information on the indicator; the opposite page contains a table and a graph providing – at a glance – the key message conveyed by the data. A dynamic link is provided for each table where readers can download the corresponding data.

The OECD Factbook is also available as a free iPhone/Smart phone application!

The OECD Health Data on CD-ROM liable to pay costs with annually actualized data has the following structure (5):

Part 1 Health Status
Part 2 Non-medical Determinants of Health
Part 3 Health Care Resources
Part 4 Health Care Utilization
Part 5 Health Care Quality Indicators
Part 6 Pharmaceutical Market
Part 7 Long term Care Resources and Utilization
Part 8 Health Expenditure and Financing
Part 9 Social Protection
Part 10 Demographic References
Part 11 Economic References
(www.oecd.org/Statistics)

The OECD edition "Health at a glance" with 64 health indicators for Member States of OECD is free of charge.

WHO Headquarter

World Health Statistics – WHOSIS Database

The WHO Geneva holds the database WHOSIS – the Statistical Information System and publish Global Health Indicators for all 194 Member States including the MDGs indicators in a special database and as a special chapter in the World Health Reports (4).

The World Health Reports refer in special topics about important aspects of health and health care. The World Health Reports of WHO have the following headlines:

2006 Working together for Health
2007 A safer Future: Global public health security in 21th Century
2008 Primary Health Care
2012 Health Systems Financing
(www.who.int/whr).

The database WHOSIS compound health indicators for 194 Member States in 10 categories which are published in the yearly World Health Statistics Report:

01 Demographic and socioeconomic indicators	15 indicators
02 Demographics	02 indicators
03 Health Equity	01 indicator
04 Health Service Coverage	20 indicators
05 Health Status	18 indicators
06 Health Systems Resources	>100 indicators
07 Morbidity and Risk Factors	13 indicators
08 Mortality	20 indicators
09 Resources, Services and Coverage	01 indicator
10 Risk Factors	> 100 indicators

World Health Statistics 2012 contains WHO's annual compilation of health-related data for all Member States, and includes a summary of the progress made towards achieving the health-related

Millennium Development Goals (MDGs) and associated targets. This year, it also included highlight summaries on the topics of non-communicable diseases, universal health coverage and civil registration coverage (www.who.int/gho/publications).

WHO Geneva has developed "Global Burden of Disease Project" and publishes periodically results for all countries of the world (www.who.int/healthinfo/global_burden_disease/en). The Global Burden of Disease project analyses a comparable assessment of mortality and loss of health due to diseases, injuries and risk factors for all regions of the world. The measure is disability-adjusted life year (DALY), that combines years of life lost due to premature mortality and years of life lost due to time lived less in full health.

Example

Indicators on reproductive health

Indicators on reproductive health developed by WHO Geneva relate to the health indicators' methodological concept of WHO, too (6). The "Reproductive Health Indicators" with Guidelines for their generation, interpretation and analysis for global monitoring have been published in 2006. The development of the shortlist of indicators for global monitoring or reproductive health followed the Millennium Summit in September of 2000, these indicators belong to the MDGs. These 17 indicators have been agreed with international agencies for monitoring the reproductive health goals. The majority of indicators belong to MDG 5. These are the following indicators:

Table 1: Shortlist of indicators for global monitoring of reproductive health (6)

1	Total Fertility Rate
2	Contraceptive prevalence
3	Maternal Mortality Ratio
4	Antenatal care coverage
5	Births attended by skilled health personnel
6	Availability of basic essential obstetric care
7	Availability of comprehensive essential obstetric care
8	Perinatal mortality rate
9	Prevalence of low birth weight
10	Prevalence of positive syphilis serology in pregnant women
11	Prevalence of anemia in women
12	Percentage of obstetric and gynecological admissions owing to abortion
13	Reported prevalence of women with genital mutilation
14	Prevalence of infertility in women
15	Reported incidence of urethritis in men
16	Prevalence of HIV infection in pregnant women
17	Knowledge of HIV-related preventive practices

WHO Regional Office for Europe, Copenhagen

After introduction of the "Health 21 – Health for All Strategy for the 21st Century", the New European policy for Health – Health 2020 is developed (7). The 6 main goals for 2014 – 2020 will be:

- Work together: WHO Member States and Regional Office to promote health and well-being, tackle the social determinants of health and health inequalities;
- Create better health: Improve the quality of life and the number of years in good health;
- Improve governance of health: Strengthen health as a driver of change for sustainable development and well-being. All key actors and decision makers should be aware of their responsibility for health;
- Set of common strategic goals: Support the development of policies and strategies that benefit health and well-being as a joint social objective;
- Accelerate knowledge-sharing and innovation: New approach to public health and person-centred health care in an ageing and multicultural society;
- Increase participation: Empower the people to be active participants in shaping health policy and to ensure the people's voice is heard in person-centred health systems.

The concept of Health 2020 is rooted in the conceptual values of health as a human right. The concept of health and well-being lie at the heart of Health 2020. The definition of health (1948) and well-being (1977) is: "Health is the state of complete physical, mental and social well-being and not merely the

absence of disease or infirmity to attain a level of health which will permit the citizen to lead a socially and economically productive life" (7).

One of the main directions of the new WHO strategy is the strong investment in prevention. With the strategies for Health targets and health indicators have been agreed. The HFA-2000 strategy set twelve global targets for health. In 1985, 38 targets for Europe were agreed, together with 65 health indicators to monitor and assess progress. The use of targets and health indicators should promote health and well-being. Up to 2020 there are in discussion 10-15 health targets. The areas for health targets will be the following seven:

1. Governance of health
2. Inequalities
3. Healthy people
4. The social determinants of health – environmental risk factors
5. Individual risk factors
6. The burden of disease, and
7. Health system performance (7).

Epidemiological and statistical data should support the Health 2020 concept and new concepts and tools for measuring health and well-being should be developed. A set of indicators was developed and agreed with an international working group [Stein C (2012) Malta (8)].

To invest in healthy people and empowering communities means to know exactly the burden of disease: gender and age groups and their health risks for non-communicable diseases, injuries, mental health, communicable diseases; vulnerable groups (i.e. migrants, Roma) families by income and education status (8).

The World Health Organization, Regional Office for Europe has developed a portal to health statistics and to detailed monitoring and assessment tools for key areas of health policy (http://www.euro.who.int/en/what-we-do/data-and-evidence/databases). WHO provided access to a broad range of information systems: from international comparisons of aggregate indicators to the results of detailed disease surveillance and the monitoring of specialized areas of health policy (9). Users can browse the information online and present and analyze it in different formats (in tables, graphs and/or maps):

- European Health for All Database (HFA-DB), which covers around 600 core indicators covering the most important chapters of a health indicator set for 53 countries. For some countries the data are not complete and not very valid. But HFA-DB is the oldest health indicator set, created after Alma-Ata Conference in 1978.
- Mortality indicator database with 67 causes of death by age and sex (HFA-MDB).
- European detailed mortality database (DMDB) with 3-digit causes of death, by gender and 5-year age groups.
- European hospital morbidity database (HMDB) based on hospital discharges by ICD-10 codes, age groups and gender.
- Centralized information system for infectious diseases (CISID) with information about communicable diseases, immunization and recent outbreaks. It is well coordinated with the European CDC in Sweden.
- European inventory of national policies for the prevention of violence and injuries.
- International inventory of documents on physical activity promotion.
- Tobacco control database.
- European Information System on Alcohol and Health (EISAH).
- European Regional Information System on Resources for the Prevention and Treatment of Substance Use Disorders (9,10).

European health reports

The European Health Reports are published every three years, the latest is from 2012. The 2012 European health report describes the overall improvements in health in the WHO European Region and their uneven distribution within and between countries. So, it defines the goals of Europe's new health policy, Health 2020 and maps the way to achieving it. One of the main goals is to chart the way to well-being. Chapter 2 deals with: European targets for health and well-being, Chapter 3 with: "The case for measuring well-being" (11). The new challenge will be to measure the well-being as a marker of progress and to find indicators for measuring well-being. For the European Health Report 2012, the available indicators have been used and allowed the overview about the health status in Europe: decline in infant mortality, in maternal mortality, in transport accident deaths and in suicides. Two important results with surprising development are:

- Cancer has replaced cardiovascular disease as the leading causes of premature death (before the age of 65) in 28 of the 53 countries in the Region, and;

- Migrants living in Europe are estimated to number 73 million (52% of whom are women) and to account for nearly 8% of the total population. Migrants are usually younger, less affluent and more likely to become ill and have less access to health services than the general population.

The European Health Report 2012 focuses particularly on well-being which forms an integral part of the new European health strategy "Health 2020", adopted by the 53 Member States in September 2012. To measure well-being and define indicators for well-being will be a challenge for the next decennium. The latest WHO report reveals unequal improvement in health in Europe and calls for measurement of wellbeing as marker of progress (11).

Example

Mental Health Indicator Set

Within the development of the ECHI indicator set the development of Mental Health Indicators was launched at a High Level Conference on Mental Health and Well-being held in June 2008 – The European Pact for Mental Health and Well-being (12). The pact covers five thematic areas in mental health:

1. Mental Health in Youth and Education
2. Preventing Suicide and Depression
3. Mental Health and Older People
4. Mental Health in Workplace Settings
5. Combating Stigma and Social Exclusion

The Health Programme 2008-2013 "Together for Health" aims to improve the level of physical and mental health and well-being of EU citizens and reduce health inequalities throughout the Community. The developed indicators cover the following areas which belong to the Mental Health Indicator Set: % of people currently depressed by gender and age groups, suicide rates by gender and age groups, satisfaction with life, calculation of wellbeing index, optimism about the future (Eurobarometer data) (Maarit Mukkala, Chris O'Sullivan, Susan Lowes; (13)).

World Health Organization (WHO)

Support for European Regions

Measurement and reporting of health conditions and actions for health improvement through "internationally agreed" indicators have been a favorite strategy of international organizations. WHO used this concept for promoting the Health for All concept strategy. The result has been a long list of indicators to be collected by the countries and to be delivered to international organizations.

During the last 10-20 years there was the development of Regions for Health Networks, starting in WHO, but now the development of Regions and their responsibility for health care is a policy driving force of European Commission. The level of regions is below the country level. In the majority of European countries the regions are responsible for health care provision. Up to now the WHO 53 European Member States deliver selected indicators about health care facilities and services as well for some health indicators like infant mortality, maternal mortality, prevalence data about mental health, cancer, COPD, etc. each year to WHO.

Regional Health Network
Common Minimum Indicator Set (CMIS)

1. Demographic and socio-economic indicators
2. Data on mortality
3. Data on prevention
4. Data on socio-economic factors
5. Data on lifestyle
6. Data on health services

During the last 10 years a set of regional (or sub-national) health indicators for the EU has been developed by the ISARE project and by EUROSTAT in order to fix objectives in health. These regional health indicators should adhere to standard definitions and methodology in order to be comparable (http://ec.europa.eu/health/indicators/other_indicators/sub_national/index_en.htm).

The regions, which are in different way defined in the Member States, are increasingly important in Europe as units for the political and administrative management of health issues. But there is up to now a problem with NUTS (Nomenclature des Unites Territoriales Statistiques) which is not usable in all countries for defining the regions.

Minimum Health Indicator Set (MHIS) for South Eastern European Countries (SEE)

Within the framework of the Stability Pact for South Eastern Europe (SEE), a Minimum Health Indicator Set (MHIS) was developed for the SEE countries which started in 2001 (16). The Minimum Health Indicator Set was based on the experience made with the Common Minimum Indicator Set (CMIS) of the Regions for Health Network, WHO-EURO, and selected from a list of 224 indicators for the WHO HFA 21 strategy 30-36 indicators.

The indicators for the pilot testing carried out in 2002 covered:

- the socio-demographic profile (percent of population aged 65+ years);
- mortality (life expectancy at birth, in years, males/females; infant mortality rate; maternal deaths, all causes; standardized death rate-SDR, circulatory system diseases, all ages, males/females; SDR malignant neo-plasms, all ages, males/females; SDR external causes injury and poison, all ages, males/females; SDR infectious and parasitic diseases, all ages, males/females);
- morbidity (number of newly diagnosed tuberculosis cases, all forms; number of decayed, missing or filled teeth at age 12);
- environment (percent of population whose homes are connected to water supply system, total; percent of population with access to hygienic on sewage disposal, total);
- health care resources indicators,(number of primary health care units per 100,000 population; number of hospital beds per 100,000 population; number of physicians per 100,000 population; number of general practitioners in PHC per 100,000 population; number of dentists per 100,000 population; number of nurses graduated per 100,000 population);
- health care utilization and costs (average length of stay, all hospitals; total health expenditures as percent of gross domestic product);
- maternal and child health (percent of infants vaccinated against diphtheria; percent of infants vaccinated against tetanus; percent of infants vaccinated against pertussis; percent of infants vaccinated against measles; percent of infants vaccinated against poliomyelitis).

This indicator set contains 30 indicators. All indicators which were included in the list, had to reflect the special situation of the South Eastern European region (SEE) (16).

After the pilot phase, 22 of the selected indicators proved to be qualified enough to reflect the health and social as well as health care situation in the PH-SEE countries. Eight indicators did not meet the quality criteria for an indicator or had to be replaced because of the poor data situation (17). The 30 final indicators are listed in Annex 2.

These 30 indicators are assigned to the following topics of the WHO Database (HFA-DB):

Demographic and socio-economic indicators	3 indicators 2 background indicators population indicators
Mortality based indicators	16 indicators
Morbidity, disability and hospital discharges	3 indicators
Lifestyles	0 indicators
Environment	0 indicators
Health care resources	4 indicators
Health care utilization and expenditure	2 indicators
Maternal and child health	2 indicators

There has been profound discussion about this health indicator list. Deciding was the availability and comparability of indicators and quality criteria. Eight indicators had to be replaced after analyzing the health situation within the PH-SEE countries and in consideration of the main topics of health policy (17).

All indicators had to meet specific criteria such as:

- relevant (regarding priorities)
- valid (regarding determinants of health)
- measurable (in quantitative and qualitative terms)
- sensitive (to changes and differences)
- comparable (inter-territorial)
- repeatable (for time series)
- affordable (in terms of relative costs)
- useful (for intervention)
- ethical (e.g. respect personal autonomy)

Definitions for all agreed 30 indicators are available in the second report of CMIS from the year 2005 (17)). The third report deals with the trend analysis for all South Eastern European Countries for the years 1990-2004 and was published in 2006 [Bardehle D, Lenz A: (18)]. It would be a work of merit to

continue the work on the Minimum Health Indicator Set in the SEE-countries and follow the development in this region. The advantage of this project was to learn the selection of a core indicator set for a special region and to analyze the development.

European Union

A design for a "Set of Community Health Indicators (ECHI I)" was developed by the European Commission under the Health Monitoring Program 1998-2003 (HMP) which contributes to the establishment of a Community health monitoring system in order to:

- Measure the health status, its determinants and trends throughout the Community;
- Facilitate the planning, monitoring and evaluation of Community programs and actions;
- Provide Member States with appropriate health information to make comparisons and support their national health policies (19,20).

The New Public Health Program (2003-2008) of the EU continued the Health Monitoring Program. It focused on three types of activities:

- To improve the quality and transparency of health information;
- To improve current abilities to respond rapidly to health threats, and;
- To find effective ways to tackle health determinants – the underlying causes of disease (19).

The strategy was supported by the European Community Health Indicators Monitoring (ECHIM) project (21). It continued the work of the previous ECHI and ECHIM projects and finished in 2012. The general objective is to consolidate and expand the ECHI Indicator system towards a sustainable health monitoring system in Europe (22,23). The focus will be on collecting and disseminating comparable health data and information based on the ECHI shortlist of 88 indicators (see annex 1) http://ec.europa.eu/health/indicators/echi/index_en.htm).

The health strategy for 2014-2020 is drafted as "The third Multi-annual program of action for health", an ambitious program to achieve innovative and sustainable health systems, promote good health of European citizens and prevent diseases and protect European citizens from cross-border threats (23). So, health determinants like smoking prevention, abuse of alcohol and obesity should be monitored and the corresponding health indicators should be available in all Member States. http://europa.eu/legislation_summaries/public_health/european_health_strategy).

The Social Investment Package – Investing in Health [EC. Brussels: (24)] establishes the role of health as part of the Europe 2020 policy framework. The improvement of data collection and using available information to underpin the improvement of the performance on health systems; in particular the collection of health data using the ECHI indicators and developing tools to better access the efficiency of health systems. There are three challenges, which concern the indicator database:

- Consider the quality of the years of life gained is more important than the crude measure of life expectancy;
- To measure the relative influence of health systems on health outcomes, including living and working conditions, income, education and lifestyle related risk factors;
- There is a time lag between policy changes and their impact on health outcomes which tends to result in "false savings" or to "increased costs".

Other challenges are the measures for reducing inequalities in health and healthcare coverage in ethnic and low-income groups and to measure the reduction of the at-risk-of poverty rate as a part of the Commission decisions on health inequalities. Up to now such indicators are not part of ECHI health indicator set.

Concerning the health information system, the Social Investment Package "Investing in Health" focuses on:

- Improving data collection;
- Using available information to underpin the improvement of the performance of health systems;
- Using the ECHI indicators and develop tools to better access the efficiency to health systems.

The assessment of the efficiency of health systems requires a refined analytical framework along three axes:

- The definition of sound, reliable indicators of health outcomes, building on ECHI indicators;
- Better understanding of the effects of health systems on health outcomes, as distinct from the impacts on health of other factors such as health determinants and lifestyles;
- Better understanding of the mechanisms, and therefore the timing, of how health policies affect health outcomes (24).

Now the "HEIDI Data Tool" started to work. HEIDI is an interactive application to present relevant and comparable information on health at European level (26). The tool is presenting a list of indicators, grouped in four sections and a fifth section under development:

- Demographic and socio-economic factors
- Health status
- Determinants of health
- Health interventions: health services
- Health determinants (under development).

Among the HEIDI database you can find the 88 ECHI indicators and among them the 40 core indicators with the highest comparability and best quality. The aim of the ECHI indicators is to provide the shortest meaningful list of indicators that gives a reliable overview on health and health system across Europe. HEIDI data tool allows query per country and per year and also per region (sub-national level). Results can be displayed in line charts, bar charts, maps or tables. The data are provided by Eurostat and other sources. Metadata description is added to each indicator.

Eurostat

Eurostat, the Statistical Office of the EU is responsible for coordinating population and health statistics as causes of death statistics, health interviews and health examination surveys and health care data groups. Population data are included in the new CRONOS database. In 2012 (1st of January) the population size of EU 27 Member States was 503.663.501 people (appsso.eurostat.ec.europa.eu). Via a public health portal, health data are presented within the EU Public Health Information Network and include databases such as EUPHIN HIEMS HEIDI (26).

One of the sources of health indicators is the European Health Interview Survey (EHIS) (30) and the Health Examination Survey. EHIS offers comprehensive data on the health status and the health-related behaviors based on household interviews. EHIS is implemented and managed by EUROSTAT and includes information from all European Union (EU) Member States. The survey is to be conducted every five years.

EHIS covers the following topics:

- Height and weight for calculation of the body mass index (BMI)
- Self-perceived health
- Activities that have been reduced because of health problems
- Long-standing illnesses or health problems
- Smoking behavior
- Alcohol consumption

With EHIS the calculation of Healthy life years and indicators about income and living conditions (EU-SILC) is possible [http://epp.eurostat.e.europa.eu/statistics_explained/index.php/Glossary:Europen_health_interview_survey_(EHIS)].

The survey methods, the content of the questionnaire and the examination protocol are available through the database and can be compared between the countries. But the results have to be interpreted, because there are ethnic, historical and other differences in living conditions between the countries, which lead in difficulties to explain differences of some indicators ie. of "perceived health" or of "limitations of daily living." With the support of EUROSTAT, the EHIS includes the candidate countries.

Within the EUROSTAT indicator database there are two parts for health data:

1. The chapter for "Public Health Data", includes:
   - Indicators on Healthy Life Years
   - Indicators on Causes of death
   - Indicators on Health care expenditure
   - Indicators on Health care resources
   - Indicators on Health care activities
   - Indicators on Health care indicators from surveys (SILC, HIS)
   - Indicators on Health status indicators from surveys (SILK, HIS, LFS)
   - Indicators on European Health Interview Survey (EHIS) – collection round 2008
2. The chapter for "Health and Safety at Work" with:
   - Indicators on health and safety at work
   - Indicators on accidents at work
   - Indicators on work-related accidents, health problems and hazardous exposure

Eurostat special Indicator Set about "Sustainable Development" contains the following groupings:

- Indicators on Socio-economic Development
- Indicators on Sustainable Consumption and Production
- Indicators on Social Inclusion
- Indicators on Demographic Changes

Indicators on Public Health
Indicators on Climate Change and Energy
Indicators on Sustainable Transport
Indicators on Natural Resources
Indicators on Global Partnership
Indicators on Good Governance
(http://epp.eurostat.ec.europa.eu/portal/page/portal/sdi/indicators/all_indicators)

The indicator Healthy Life Years as the leading EU Health Indicator is included in the "Structural Indicator Set" of Eurostat besides a lot of social indicators, too.
Another indicator set for all EU-Member States, hold by EUROSTAT are 31 long-term Health Indicators within the topic "Population and social conditions". These indicators belong to following topics:

Public Health / Heath Status	4 indicators
Health Care Statistics	7 indicators
Causes of Death Statistics	14 indicators
Health and Safety at Work	6 indicators

Additionally, indicators on diagnoses specific morbidity statistics have been added, starting in 2007.

Data and indicator delivery from European Countries to International Organizations

The countries have to deliver a lot of data and indicators to international organizations. This process has to be organized very carefully and the responsibility has to be clear, otherwise you will find deviations between countries' data and international published data. Which organizations need data from the countries? We will list up only a selection of requirements for health data:

- HFA-DB with around 600 indicators
- Eurostat databases and indicator sets of which: 17 health indicators for "sustainable development indicators" and 31 health indicators for "long-term indicators";
- 400 ECHI indicators, of which 88 core indicators: shortlist;
- Millennium Development Goals with 18 health indicators.

For WHO and Eurostat mortality databases, the countries have to deliver disaggregated data by gender, diagnoses, so called shortlists with 65-67 diagnoses groups of even more. The international data delivery and the quality assurance for these data are big challenges for countries. Additionally, you will have a country-specific health indicator set with around 100 health indicators and you have to deliver indicators to NGOs, to UNICEF and some sponsors. A lot of indicators are the same in the listed up databases, but for some similar indicators you have to calculate with different methods. And you need the overview in your country, who is delivering data, what is the time-schedule for data delivery and do you have storage for the delivered data. Additionally you will need skilled personnel for this task. Delivery of so called vital statistical data is in the responsibility of the country's Statistical Offices, but some other data like incidence and prevalence data, health care facility and service data and manpower data are in the responsibility of the Ministry of Health and the National Public Health Institute.

Up to now the 53 European Member States deliver selected indicators about health care facilities and services as well for some health indicators like infant mortality, maternal mortality, prevalence data about mental health, cancer, COPD etc. Every year, other data are directed to a common portal to WHO, Eurostat and OECD. These data are used by EU, WHO and OECD and must not be delivered three times.

Indicator classification and evaluation methods

Definitions of "Health Indicators" (www.who.deficrit.htm)

Indicators are markers of the health status, healthcare system performance or availability of resources, defined in a way to allow the monitoring of objectives, targets and performance. Thus they cannot be confused with objectives and targets. Objectives are statements aiming to improve health or to reduce the frequency of certain diseases, expressed in a quantitative manner, within a given time frame. Targets are usually expressions of the desired service performance, for example, output or coverage, desired to be achieved at some time point in the future. Indicators are defined as variables able to measure the changes in the level of health target achievement i.e. Health for All (HFA) targets.

The European Commission (Public Health) defined indicators in the following way:
"An indicator is a quantitative or qualitative measure of how close we are to achieving a set goal (policy outcome). They help us analyze and compare performance across population groups or geographic areas, and can be useful for determining policy priorities. Health indicators based on reliable, comparable data are essential for designing strategies and policies to improve the health of Europeans, and then monitoring their implementation. (http://ec.europa.eu/health/indicators/policy/index_en.htm)

Another EC definition of indicators from 2005 expressing the interrelation between data sets, health policy and health indicators is the following:

“Indicators are on the crossroad of policy questions and data sets, they have to combine some objectives.

The health information strategy aims to set out an overall approach for generating, disseminating and applying information, based on the priorities outlined in the health strategy, and create an organizational structure and methods.

Indicators are used for health monitoring and health surveillance

Health monitoring is defined as the maintenance or regular checking of ongoing activities or programs with respect to predefined objectives. The purpose is to record what the system is actually achieving at present and to detect possible deviations from the decided course of action.

Surveillance refers to the ongoing observation of the health status of a population and the factors that may affect it, and its purpose consists in detecting possible changes at an early stage and initiating appropriate action (21).

Types of indicators

There are three types of HFA indicators which are defined within the HFA 21 strategy for health indicators (27,28). Definitions and criteria are:

- Outcome (health status or death);
- Process (health care delivery and management, including resources, input indicators, health care performance indicators);
- Determinants (e.g. behavioral factors and public knowledge).

All HFA 21 indicators (28) can be used to measure progress towards established targets and goals, including the monitoring of changes in the health status of the population. Most of them can be used to monitor service performance at the facility, at district and national levels.

Generic indicators are broadly defined areas of measurements linked to specific parts of the HFA policy framework (HFA targets) and traditionally constitute an integral part of the HFA policy document. Operational indicators are precisely defined numerical data items as recorded in the HFA statistical database (27,28).

An indicator can be defined at the generic level, e.g. „smoking behavior”, or in an operational manner, e.g. “% of women in x age group, x smoking between y and z cigarettes per day”. Operational indicators are always expressed in a numerical way, calculated from primary data in a more or less complex manner. An example of a complex calculation is „life expectancy at birth”, which is calculated from a large set of age-specific mortality data. Indicators are usually numerical (ratios, proportions, rates), although they can also be qualitative (e.g. existence or absence of a sign, event, etc. that has been shown to be important).

Quality criteria for health indicators

With regard to the selection of indicators, the following prerequisites are necessary:

- The actual selection and definition of indicators within a specific public health area should be based on scientific principles.
- Indicators (and underlying data) should meet a number of methodological and quality criteria concerning e.g. quality, validity, sensitivity and comparability.
- The probability of changing policy priorities/interests calls for a high degree of flexibility, made possible through current electronic database systems.
- The selection of indicators should be based on existing and comparable data sets for which regular monitoring is feasible, but should also indicate data needs and development areas (27,28).

The quality of indicators will be measured according to the following four criteria:

1. Validity: The indicator is a true expression of the phenomena it is measuring;
2. Objectivity: The indicator is able to provide the same result if measured by different people under similar circumstances;
3. Sensitivity: The indicator is capable of reflecting changes in the phenomena of interest;
4. Specificity: The indicator reflects changes in the specific phenomena of interest only.

Additionally, the following criteria are relevant for the use of an indicator and the methodology employed to collect the data:

- The data required for the indicator are useful for case management or taking action in the community for the staff who originally recorded the data, or the service unit from which the data originated.
- It should be feasible to obtain the data needed for each indicator and that these data should be generated, as far as possible, through routine service processes or through easily and rapidly executable surveys.
- Indicators should be simple and understandable, measuring a health condition or aspect of service.

- The indicator and the process of collecting and processing the relevant data are ethical (27,28).

Health indicators serve several purposes:

1. Indicators are an important tool of for health policy formulation and implementation.
2. Indicators are used to track progress, i.e. they are used for monitoring and evaluating the health situation with respect to specified objectives.
3. Indicators can provide yardsticks / benchmarks whereby countries can compare their own progress with that of other countries, especially those at similar levels of socio-economic development.
4. Indicators cannot be measured at present because no adequate information is in place; they are nevertheless adopted for use because they point to what needs to be done (guidance for action, including information systems' development).
5. Indicators have a communication and coordination function: for example, when decided in a proper consultation process they constitute an important message to the community about agreed priorities (27,28).

Main categories of an indicator set

The following main categories of a set of Community Health Indicators (ECHI Indicator set) were proposed (21):

1. Demographic and socio-economic factors
- 1.1. Population
- 1.2. Socio-economic factors

2. Health status
- 2.1. Life expectancy
- 2.2. Infant / perinatal mortality
- 2.3. Mortality, disease-specific
- 2.4. Morbidity, disease-specific, communicable / non communicable
- 2.5. Injuries
- 2.6. Self-perceived health
- 2.7. Activity limitations
- 2.8. Health expectancy / Healthy life years

3. Determinants of health
- 3.1. Personal and biological factors / BMI / RR
- 3.2. Health behaviors / life style / smoking/ alcohol / illicit drugs
- 3.3. Eating behaviors
- 3.4. Physical activities
- 3.5. Work-related risks
- 3.6. Social support

4. Health interventions / Health services
- 4.1. Prevention, health protection and health promotion
  vaccination coverage, screening participation
- 4.2. Health care resources and manpower
- 4.3. Health care utilization / inpatient and outpatient
- 4.4. Health expenditures and financing
  insurance coverage
- 4.5. Health care quality / Performance
  waiting times, wound infections, disease control programs

5. Health interventions / Health promotion
- 5.1. Health policy programs (smoking, nutrition, lifestyle)
- 5.2. Integrated programs in settings (schools, workplace)

This structure corresponds to the Development Model for classification of health indicators published by Bauer G et al. and sponsored by European Commission (26).

The proposed structure is the following:
1. Demographic and socio-economic factors.
2. Health status.
3. Determinants of health.
4. Health systems incl. health promotion.

This structure was developed within the first Health Monitoring Program under P. Kramers (Bauer G).

Corresponding to the ECHI concept the indicators could be divided into the following categories:

1. Cockpit information, now they are renamed into core indicators which are summarized in the shortlist of ECHI indicators: to have a quick view on the major trends in public health, including recent relevant signals, for medium or long-term policy strategies; the shortlist is a subset of the long-list.
2. EU priority list: to follow developments for specific EU policy areas or targets, programs or projects; such health indicators we find in the "Sustainable Development Indicator Set" and in the "Structural Indicator Set" in the EUROSTAT databases.
3. User-windows are subsets of indicators selected from a specific perspective, i.e. health indicators for health promotion, indicators for health and services for mother and child (reproductive health), indicators for mental health.

(Annex 5 to the ECHI-2 report of June 20, 2005. The ECHI comprehensive indicator list (25). (long list, version of July 7, 2005).
www.ec.europa.eu/health/ph_information/indicators

4. Long-list of indicators, i.e. the comprehensive ECHI Indicator list (25).

The operational indicators of the Health 21 Strategy

The HFA indicators are designed to be an integral part of the HFA policy framework adopted by the Member States of the WHO Region for Europe. They serve as the main tool for monitoring progress toward attainment of the HFA targets and for assessing the health situation in European countries in general (28).

The indicators are formulated in the health policy framework of WHO at the generic, conceptual level. Generic indicators are loosely defined areas of measurements linked to specific parts of the HFA policy framework (HFA targets) and traditionally constitute an integral part of the HFA policy document. The list with generic indicators for HEALTH21was approved with Member States in 1998. The main change between the 1991 generic indicators and the HEALTH21 proposal from 1998/1999 was a reduction in the total number of generic HFA indicators from 112 to 59. 50 indicators have been used from HFA 1991 indicators, but 9 generic indicators have been adopted.

The next step was the definition of operational indicators for HEALTH21. Operational indicators are precisely defined numerical data items as they are kept in the HFA statistical database. WHO Regional Office for Europe defined with an expert group from 16 countries, including European Commission and OECD the demands on operational health indicators to monitor the Health 21 strategy [The Hague, 2-3 March, 2000; (28)].

Operational indicators are divided into the following groups [WHO Copenhagen, The Hague 2000, Annex 3: (28)]**:**

- Indicators on mortality, by gender, Age standardization, secondary data sources
- Indicators on morbidity (secondary data sources and epidemiological register data)
- Indicators on disability (new invalidity, social benefits, long-term disability, special registers or surveys)
- Indicators on maternal and child health (secondary data and epidemiological data, survey data)
- Other health status indicators (4 indicators, of which DALY, self-assessment of health)
- Indicators on lifestyle (smoking, alcohol, drugs, food, BMI)
- Indicators on environment (living area per person)
- Indicators on health care resources (hospitals, primary health care, manpower)
- Indicators on health care utilization
- Indicators on quality of care (autopsy rate in hospital deaths, surgical wound infection rate, renal failure, blood transfusion)
- Indicators on health expenditure (total health expenditure, public health expenditure, salaries)
- Demographic and socio-economic indicators (mid-year population, live births, fertility rate, employment rate, literacy rate, education level.

Data for indicators are being collected from various sources (HFA 21) (28). The main information sources are (29):

- Comprehensive statistical records already established for health or other purposes;
- Ad hoc investigation or surveillance systems within the health services, and;
- Population surveys.

All efforts are made to use information from available sources to avoid duplicating requests to countries.

Data Source: Health Interview Surveys (HIS) and Health Examination Surveys (HES)

In 1988, 1990 and 1992 the WHO Regional Office for Europe and Statistics Netherlands organized consultations to develop common methods and instruments for a health interview survey at the European level [HIS WHO 1996; (29)]. The objective was that this health interview survey should be used in all

countries in order to achieve better international comparability and enhance the value and use of survey results. In order to measure the improvements in health through programs, HFA-2000 needs quantitative and qualitative indicators for the 36 health targets and 65 health indicator sets which have been agreed with all Member States. For some part of the indicators, especially for qualitative indicators, the countries need indicators from Population surveys. They include mainly the areas of [WHO HIS 1996: (29)]:

- Self-perception of health status and the indicators related to disablement;
- Lifestyle-related indicators (smoking, alcohol consumption patterns, breast feeding, physical activity and indicators related to health promotion;
- (in some countries) indicators related to environmental health (water, sanitation, housing);
- Those aspects of health service provision and use where individual response is a major factor (e.g. family planning).

The European Union continued the work on methodological background of HIS and HES for all Member States and conduct systematic health interview surveys for EU-Member States every three years [Aromaa A 2003: (30)]. The Health Interview Survey and Combinations with Health Examination Survey are central components of a comprehensive health monitoring system.

HISs deliver valuable information on health status, illnesses, lifestyles, functional capacity and use of health care services. For data on perceived health, symptoms and health-related behavior the Member States need Health Interview Surveys (Aromaa A). In addition, clinical measurement is needed to obtain valid information on many chronic conditions, functional limitations and disabilities, and on several key health determinants.

Recommended instruments for health interview surveys are:

Demographic and socioeconomic factors
01. Education and / or employment status

Health Status
02. Self assessed / Perceived health
03. Long-standing illness / chronic conditions / disability
04. Limitations of activities of daily living, personal
05. General mental health

Personal factors
06. Body weight and height

Life style factors
07. Smoking
08. Alcohol use and abuse
09. Physical activity

Working conditions, Housing conditions, Prevention
10. Contraception
11. Vaccinations

Health and social services
12. Use of services, GP
13. Medication
14. Hospitalization
[Aromaa A, 2003, Table 4: Analysis of 60 surveys during 1998-2002, topics covered in interviews: (30)].

Methods of comparison and benchmarking

The application of statistical methods will be the subject of other parts of the curriculum. To complete the establishment of indicator sets and use of health indicators, it has to be mentioned that comparability must be guaranteed with the help of the following methods. The use of statistical methods for comparing data of different regions includes:

- Age standardization incl. calculation of confidence intervals;
- Significance check-ups;
- Unified and agreed calculation for composite indicators, i.e. healthy life expectancy;
- Definitions of the included regions concerning the application of further statistical methods;
- Calculations such as “PYLL: Person years life lost”;
- Calculations for time trends, and;
- Meta-database description of the data used including definitions.

The use of health indicators for health reporting

Currently, various methods are used for health reporting:

- Indicator-based health reporting

On the basis of a well-defined indicator set, periodic health reporting is done to follow the indicators and trends. Changes of in the indicator level are analyzed and described within the different chapters of periodic health reporting.

• Indicator sets and their use for health reporting

For writing health reports with the help of experts or for special topics (e.g. women's health) a part of the indicators sets can probably be used, but usually not the complete indicator set. The advantage consists in the flexibility of the reporting, the disadvantage is the discontinuity of a frame for reporting such as "Health situation in South Eastern Europe". However, within the Stability Pact a report based on the Minimum Indicator Set for South East European countries was produced in 2003 (8) and can serve as a model for future similar reports aiming to support the decision-making process in the area and to track progress of these countries towards the goals of integration in the European Union.

• Health targets, health indicators and health reporting

WHO prefers health reporting on the basis of health targets. The advantage is the good tracking of the targets. A good example is the UK model or the "Healthy People" strategy of the United States (www.health.gov/healthypeople). The disadvantage lies in the time-consuming process of formulating common targets for several countries. Also the establishment of an indicator set with benchmarking criteria based on health targets takes a lot of time and is a difficult undertaking. Some targets may change in the course of the years and so you have to change your indicator set, too. Here WHO has made some experience. Thus the indicators of based on the new strategy HFA 21 are more "generic" and less "operational".

A review on health target setting in 18 European countries (15) demonstrated that Health for All strategy has influenced the health policy of almost all of the 18 countries.

Exercises

Health Indicators and Health Reporting

Task 1: After being familiar with the HFA software, students are asked to select a set of relevant indicators from this database and to prepare a report describing the situation from a certain country/region for the purpose of priority setting. Time: 120 minutes.

Task 2: Students are asked to search actual information for the Minimum Health Indicator Set (see: annex 2) using the HFA-DB and make comparisons between SEE countries (e.g. in life expectancy at birth, infant mortality rate and SDR due to different causes) and try to find possible explanations.

Task 3: Review existing national data sources (available in your country) and look for available indicators also describing the local levels (e.g. district, country, etc.) and make geographical comparisons. Commonly, reports or databases are reported by National Statistical Institutes/Bureaus and Institutes of Public Health.

References

1. United Nations. Millennium Development Goals. United Nations, Department of Economic and Social Affaires/StatisticsDivision/Populationstatistics.www.un.org/millenniumgoalshttp://www.un.org/en/development/desa/index.html (Accessed 21 March 2013).
2. United Nations (2013) UN Millennium Project 2000-2015. Millennium Development Goals. http://www.unmillenniumproject.org/ (Accessd 05 March 2013).
3. United Nations (2012): The Millennium Development Goals Report 2012. http://mdgs.un.org/unsd.mdg (Accessed 21 March 2013).
4. WHO Headquarter (2013) WHO Statistical Information System (WHOSIS). www.who.int/whosis/data (Accessed 09 April 2013).
5. Organization for Economic Co-Operation and Development (OECD) (2013): www.oecd.org/statistics; www.oecd.org/documents/explanatory_notes (Accessed 09 April 2013).
6. WHO Headquarter. Department of Reproductive Health and Research (2006) Reproductive Health Indicators for Global Monitoring. 2nd report. Unedited. www.who.org/reproductivehealth/publications/monitoring (Accessed 21 March 2013).
7. World Health Organization. Regional Committee for Europe (2011) The new European Policy for Health – Health 2020. Vision. Values, Main Directions and Approaches. (EUR/RC61/9). www.euro.who.int/_data/assets/pdf_file/ (Accessed 05 March 2013).

8. Stein C (2012) Targets and Indicators for Health 2020. Progress Reports and Next Steps. WHO Regional Office for Europe, Division of Information, Evidence, Research and Innovation. Presentation of the EUPHA Congress Malta, 10-13 September, 2012.
9. WHO Regional Office for Europe Copenhagen and London, 13 March 2013. Information for the media. www.euro.who.int/en/what-we-publish/information-for-the-media/sections (Accesed 22 March 2013).
10. World Health Organization. Regional Committee for Europe (2013) Data and Evidence Databases. http:// www.euro.who.int/en/what-we-do/data-and-evidence/databases (Accessed 05 March 2013).
11. WHO Regional Office for Europe (2013: European Health Report 2012. www.euro.who.int/what-we-do/the_european_health_report_2012. (Accessed 09 April 2013).
12. O'Sullivan Chris (2010): Mental Health Indicators and Data in EU Member States. Preventing Suicide and Depression. Establishment of a set of mental health indicators for the European Union. Scottish Development Centre for Mental Health Edinburgh, in collaboration with DG Health and Consumers (2003- 2008) (Accessed 21 March 2013).
13. Mukkala M, O`Sullivan Ch, Lowes S (2013) Mental Health Indicators and Data in EU Member States – Preventing Suicide and Depression) Scottish Development Centre for Mental Health, Edinburgh.
Supported by DG SANCO. (Accessed 15 March 2013).
14. WHO Europe. Regions for Health Network (2000) Common Minimum Indicator Set – Minimum essential information for the HFA 21 policy on regional level with 30 indicators www.lzg.gc.nrw.de/ (Accessed 09 April 2013).
15. Fuhr A, Bardehle D, Monch M (2001) Common Minimum Indicator Set. WHO Europe. Regions for Health Network. Final Report 2000. Wissenschaftliche Reihe des LOGD Band 15.
16. Zalatel-Kragelj L, Bardehle D, Burazeri G, Donev D, Laaser U. Minimum Health Indicator Set (MHIS) for PH-SEE countries. Final Report. Bielefeld: PH-SEE Programme for Training and Research in Public Health. Bielefeld; 2003. Available from URL http://www.snz.hr/ph-see/docs.html.
17. Bardehle D, Lenz A. Minimum Health Indicator Set for PH-SEE Countries. Second Report 2005. Stability Pact for South Eastern Europe. Public Health Collaboration in SEE. Institute of Public Health North Rhine-Westphalia, Germany. Bielefeld 2005.
18. Bardehle D, Lenz A. Minimum Health Indicator Set (MHIS) for PH-SEE Countries. Trend Analysis for 1990-2004. Third Report, PH-SEE Programmes for Training and research in Public Health. State Institute of Public Health of North Rhine-Westphalia. Bielefeld, 2006.
19. European Union, Health and Consumer Directorate General. New Public Health Programme 2003-
20. 2008. Available from URL: http://europa.eu.int/comm/health/ph_programme.
21. European Union (2007) Commission White Paper of 23 October 2007: Together for Health. A Strategic Approach for the EU 2008-2013 (COM (2007) 630).
http://europa.eu/legislation_summaries/public_health/european_health_strategies (Accessed 05 March 2013)
22. European Community (2008) European Health Indicators Development and Initial Implementation. Final Report of the ECHIM Project. National Public Health Institute Helsinki, 2008. www.echim.org/docs/echim_final_report (Accessed 26 March 2013).
23. European Commission (2013) Indicators: European Community Health Indicators (ECHI). Complete list with 88 indicators and 40 core indicators.
http:// ec.europa.eu/health/indicators/echi/index_en.htm. (Accessed 05 March 2013).
24. European Union (2011) Proposal for a Regulation of the European Parliament and of the Council of 9 Nov 2011 on establishing a Health for Growth Program, the third multiannual program of EU action in the field of health for the period 2014-2020 (COM(2011) 709).
http://europa.eu/legislation_summaires/public_health/european_health_strategies
(Accessed 05 March 2013).
25. European Commission (2013) Social Investment Package. Commission Staff Working Document. Investing in Health. Brussels, 20.2.2013 (SWD (2013) 43 final).
http://ec.europa.eu/health/strategy/docs (Accessed 05 March 2013).
26. European Commission. DG V (2005) Annex 5 to the ECHI-2 report of June 20, 2005. The ECHI comprehensive indicator list. (long list, version of July 7, 2005).
www.ec.europa.eu/health/ph_information/indicators (Accessed 04 April 2013).

27. EUROSTAT (2013) HEIDI Data Tool. Health Indicators. http://ec.europa.eu/health/indicators (Accessed 09 April 2013).
28. Bauer G, Davies J K, Pelikan J (2006) The EUPHID Health Development for the classification of public health indicators. Health Promotion International (2006) 21 (2), pp. 153-15.
29. WHO Regional Office for Europe (2000): WHO HFA Indicators for the New Public Health Policy in Europe. Report on a WHO Expert Group Meeting. The Hague, 2-3 March, 2000. EUR/00/5018702. www.euro.who.int (Accessed 05 April 2013).
30. WHO Regional Office for Europe (1996): Editors: de Bruin A, Picavet H S J, Nossikov a. Health Interview surveys: Towards international harmonization of methods and instruments: who REGIONAL PUBLICATIONS: European series; No. 58. www.euro.who.int/_data/assets/pdf_file (Accessed 05 April 2013).
31. Aromaa A, Koponen P et al.: Evaluation of Health Interview Surveys and Health Examination Surveys in the European Union. European Journal of Public Health (2003) 13 (supplement): 67-72.

Annex 1

European Commission
Public health Indicators
ECHI – list of indicators (88)
http://ec.europa.eu/health/indicators/echi/list/index_en.htm (Accessed 05 March 2013)

The already implemented indicators in the HEIDI data tool are marked (I)
Indicators under development are marked with (D)
Other useful indicators additionally to ECHI indicator set are marked with the symbol of DG SANCO
If possible and available stratification by gender and age and by socioeconomic or regional level is provided.

Demographic and socio-economic situation

1.	Population by sex/age	(I)
2.	Birth rate, crude	(I)
3.	Mother's age distribution	(I)
4.	Total fertility rate	(I)
5.	Population projections	(I)
6.	Population by education	(I)
7.	Population by occupation	(D)
8.	Total unemployment	(I)
9(a)	Population below poverty line	(I)
9(b)	Income inequality	(I)

Health status

10.	Life expectancy	(I)
	Life expectancy by educational attainment	(I)
11.	Infant mortality	(I)
12.	Perinatal mortality	(D)
	European health indicator on Perinatal mortality rate	DG Sanco
13.	Disease-specific mortality	(I)
14.	Drug-related deaths	(I)
15.	Smoking-related deaths	(D)
16.	Alcohol-related deaths	(D)
17.	Excess mortality by heat waves	(D)
18.	Selected communicable diseases	(I)
	European health indicator on communicable diseases	DG Sanco
19.	HIV / AIDS	(D)
	European health indicator on HIV / AIDS	DG Sanco
20.	Cancer incidence	(D)
21(a).	Diabetes: self-reported prevalence	(I)
21(b).	Diabetes: register-based prevalence	(D)
22.	Dementia	(D)
	European health indicator on dementia	DG Sanco
23(a).	Depression: self-reported prevalence	(I)
23(b).	Depression: register-based prevalence	(D)
24.	Acute myocardial infarction (AMI)	(D)

25.	Stroke	(D)
26(a).	Asthma: self-reported prevalence	(I)
26(b).	Asthma: register-based prevalence	(D)
27(a).	Chronic obstructive pulmonary disease (COPD): self-reported prevalence	(I)
27(b).	Chronic obstructive pulmonary disease (COPD): register-based prevalence	(D)
28.	Low birth weight	(I)
29(a).	Injuries: home, leisure, school: self-reported incidence	(D)
29(b).	Injuries: home, leisure, school: register-based incidence	(D)
30(a).	Injuries: road traffic: self-reported incidence	(D)
30(b).	Injuries: road traffic: self-reported incidence	(D)
	European health indicator on road accidents	DG Sanco
31.	Injuries: workplace	(I)
32.	Suicide attempt	(D)
33.	Self-perceived health	(I)
34.	Self-reported chronic morbidity	(I)
35.	Long-term activity limitations	(I)
36.	Physical and sensory functional limitations	(D)
37.	General muskulo-skeletal pain	(D)
38.	Psychological distress	(D)
39.	Psychological well-being	(D)
40.	Healthy expectancy: Healthy life years (HLY)	(I)
41.	Health Expectancy, others	(I)
Determinants of health		
42.	Body mass index	(I)
43.	Blood pressure	(I)
44.	Regular smokers	(I)
45.	Pregnant women smoking	(D)
46.	Total alcohol consumption	(I)
47.	Hazardous alcohol consumption	(D)
48.	Use of illicit drugs	(I)
49.	Consumption of fruit	(I)
50.	Consumption of vegetables	(I)
51.	Breastfeeding	(D)
52.	Physical activity	(I)
53.	Work-related health risks	(D)
54.	Social support	(D)
55.	PM10 (particulate matter) exposure	(I)
Health interventions: health services		
56.	Vaccination coverage in children	(I)
57.	Influenza vaccination rate in elderly	(I)
58.	Breast cancer screening	(I)
59.	Cervical cancer screening	(I)
60.	Colon cancer screening	(I)
61.	Timing of first antenatal visits among pregnant women	(D)
62.	Hospital beds	(I)
63.	Physicians employed	(I)
64.	Nurses employed	(I)
	European health indicator on other health professionals	DG Sanco
65.	Mobility of professionals	(D)
66.	Medical technologies: MRI units and CT scans	(I)
67.	Hospital in-patient discharges, limited diagnoses	(I)
68.	Hospital day cases, limited diagnosis	(D)
69.	Hospital day cases / in-patient discharge ratio, limited diagnoses	(D)
70.	Average length of stay (ALOS), limited diagnoses	(I)

71(a).	General practitioner (GP) utilization; self-reported visits	(D)
71(b).	General practitioner (GP) utilization; registered visits	(D)
72(a).	Selected outpatient visits: self-reported visits	(D)
	European health indicator on self-reported visits to a dentist or orthodontist	DG Sanco
72(b).	Selected outpatient visits: registered visits	(D)
73.	Selected surgeries	(I)
74.	Medicine use, selected groups	(I)
75.	Patient mobility	(D)
76.	Insurance coverage	(I)
77.	Expenditures on health	(D)
78.	Survival rates cancer	(D)
79.	30-day in-hospital case-fatality of AMI and stroke	(D)
80.	Equity of access to health care services	(I)
	European health indicator on equity of access to dental care services	DG Sanco
81.	Waiting times for elective surgeries	(D)
82.	Surgical wound infections	(D)
83.	Cancer treatment quality	(D)
84.	Diabetes control	(D)
Health interventions: health promotion		
85.	Policies on environmental tobacco smoke (ETS) exposure	(D)
86.	Policies on healthy nutrition	(D)
87.	Policies and practices on healthy lifestyles	(D)
88.	Integrated programs in setting, including workplace, schools, hospitals	(D)

http://ec.europa.eu/health/indicators/echi/list/index_en.htm (downloaded: 05.03.2013)

Annex 2

List of the Minimum Health Indicator Set (MHIS) for South-Eastern European Countries

(Stability Pact for South Eastern Europe. Public Health Collaboration in South Eastern Europe (Editor) (Layout and Editing: Prof. Dr. Doris Bardehle, Institute of Public Health North Rhine-Westphalia (Germany) Alexander Lenz, University Bielefeld. Minimum Health Indicator Set for PH-SEE Countries. Second Report. Bielefeld 2005.

Indicator	Description	HFA-DB code
01	% of population aged 65+	0030 999902
02	Live births per 1,000 of population	0060 998004
03	Unemployment rate in %	0200 020501
04	Life expectancy at birth, in years, male	1011 060101
05	Life expectancy at birth, in years, female	1012 060101
06	Infant deaths per 1,000 live births	1110 070100
07	Perinatal death per 1,000 births	1170 070403
08	Maternal deaths per 100,000 live births	1210 080100
09	Maternal deaths, abortion per 100,000 live births	1211 080101
10	SDR, all causes, all ages per 100,000, male	1811 990102
11	SDR, all causes, all ages per 100,000, female	1812 990102
12	SDR, diseases of circulatory system, all ages per 100,000, male	1321 090102
13	SDR, diseases of circulatory system, all ages per 100,000 female	1322 090102
14	SDR, malignant neoplasms, all ages per 100,000, male	1521 100102
15	SDR, malignant neoplasms, all ages per 100,000, female	1522 100102
16	SDR, external causes injury and poison, all ages, per 100,000, male	1721 110102
17	SDR, external causes injury and poison, all ages, per 100,000, female	1722 110102

18	SDR, infectious and parasitic diseases, all ages, per 100,000, male	1821 993002
19	SDR, infectious and parasitic diseases, all ages, per 100,000, female	1822 993002
20	Tuberculosis incidence per 100,000 of population	2010 040301
21	Measles incidence per 100,000 of population	2080 050111
22	Diphtheria incidence per 100,000 of population	2100 050113
23	Hospital beds per 100,000 of population	5050 270205
24	Physicians per 100,000 of population	5250 270201
25	General practitioners (PP) per 100,000 of population	5290 992733
26	Dentists (PP) per 100,000 of population	5300 270203
27	Average length of stay, all hospitals, in days	6100 992901
28	Total health expenditure as % of gross domestic product (GDP)	6710 340102
29	% of infants vaccinated against diphtheria in 1st year of life	7160 280101
30	% of infants vaccinated against poliomyelitis in 1st year of life	7200 280105

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Measuring burden of disease: Disability-Adjusted Life Years (DALY)**
Module: 2.30	ECTS (suggested): 0.4
Author(s), degrees, institution(s)	Doncho Donev, MD, PhD, Professor Institute of Social Medicine, Faculty of Medicine, University "Ss Cyril and Methodius", Skopje, Republic of Macedonia Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Vesna Bjegovic, MD, PhD, Professor Institute of Social Medicine and Centre School of Public Health, School of Medicine, University of Belgrade, Serbia Genc Burazeri, MD, PhD Department of International Health, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands
Address for correspondence	Doncho Donev Institute of Social Medicine, Faculty of Medicine, 50 Divizia 6, MKD-1000 Skopje, Republic of Macedonia E-mail: dmdonev@gmail.com
Keywords	Burden of disease, DALYs, summary measures of the population health.
Learning objectives	After completing this module students and public health professionals should be able to: • understand the general approach in measuring the burden of diseases (BoD); • understand the concept of disability-adjusted life year (DALY); • perform DALY calculations using a basic procedure on an individual basis, as well as on population basis.
Abstract	The BoD concept is used in examining health issues internationally from the perspective of determining cost-effective interventions and priority setting for resource allocation. The DALY has emerged as a measure of the BoD in early 1990-ties. DALYs for a disease or health condition are calculated as the sum of the Years of Life Lost (YLL) due to premature mortality and the Years Lost due to Disability (YLD) of the health condition of less than full health.
Teaching methods	An introductory lecture gives the students insight into DALY concept and calculation procedure. In continuation they apply in small groups calculation according to examples and case study provided, discuss the results and problems.
Specific recommendations for teachers	Work under teacher supervision/individual students' work proportion: 30%/70%; Facilities: a lecture room, a computer room; Equipment: computers, LCD projection, access to the Internet, access to the bibliographic data-bases; Training materials: recommended readings or other related readings; Target audience: master degree students according to Bologna scheme.
Assessment of students	A multiple-choice test and practical calculation of a given case study.

MEASURING BURDEN OF DISEASE: DISABILITY ADJUSTED LIFE YEARS (DALY)

Doncho Donev, Lijana Zaletel-Kragelj, Vesna Bjegovic, Genc Burazeri

Theoretical background

Burden of disease

Various diseases that people suffer from put different amount of disease-burden on populations. In last decades this burden is increasingly measured across the nations with intention to compare it. We could define burden of disease (BoD) as the burden that a particular disease process has in a particular area as measured by cost, morbidity, and mortality. It is quantified by the so-called summary measures of population health.

Summary measures of population health

Summary measures of population health are measures that combine information on mortality and non-fatal health outcomes to represent the health of a particular population as a single number. Over the past 30 years or so, several indicators have been developed to adjust mortality to reflect the impact of morbidity or disability. These measures fall into two basic categories, health expectancies and health gaps (1-3).

1. **Health expectancies**
   Health expectancies measure years of life gained or years of improved quality of life. In this group of measures, among others, the following measures are classified:
   - active life expectancy (ALE)
   - disability-free life expectancy (DFLE)
   - disability-adjusted life expectancy (DALE)
   - healthy adjusted life expectancy (HALE)
   - quality adjusted life expectancy (QALE)
2. **Health gaps**
   Health gaps measure lost years of full health in comparison with some 'ideal' health status or accepted standard. In this group of measures among others, the following measures (indicators) are classified:
   - potential years of life lost (PYLL)
   - healthy years of life lost (HYLL)
   - quality adjusted life years (QALY)
   - disability adjusted life years (DALY)

Both approaches use time and multiply number of years lived (or, not lived in case of premature death) by the "quality" of those years. The process of adjustment of the years of healthy life lived is called "quality adjustment" (expressed as QALYs), and the process of adjustment of the years of healthy life lost is called "disability adjustment" (expressed as DALYs) (4,5). It means that QALYs represent a gain which should be maximised, and DALYs represent a loss which should be minimised. In the QALY approach the quality is weighted (sometimes called "utility", as it is the case of cost-utility analyses) on a scale from 1 indicating perfect health and the highest quality of life, to 0 indicating no quality of life and is synonymous to death. In the DALY approach the scale goes in opposite way: a disability weighted zero indicates perfect health (no disability), and weighted 1 indicates death. The disability weighting is the most difficult and controversial part of the DALY approach (5). A typology of summary measures of population health is presented in Figure 1.

There are two lines (upper and lower) and three areas (A, B and C) (Figure 1). The upper line is the survivorship curve from a hypothetical life table population. The lower curve is a hypothetical curve of survivors to each "age x" in optimal health. Area A represents time lived in optimal health, area B time lived in suboptimal health, and area C time lost due to mortality. Total life expectancy at birth is given by the area under the upper curve (Equation 1):

$$LE = A + B \qquad \textbf{Equation 1.}$$

LE = total life expectancy at birth

A = time lived in optimal health

B = time lived in suboptimal health

Figure 1. A typology of summary measures. Adapted from Murray CJL and Lopez AD (1)

LEGEND: A=time lived in optimal health, B=time lived in suboptimal health, C=time lost due to mortality.

Health expectancies are population indicators that estimate the average time that a person could expect to live in a defined state of health. In terms of Figure 1, health expectancy is given by following equation (Equation 2):

$$HE = A + f(B) \qquad \textbf{Equation 2.}$$

HE = health expectancy
A = time lived in optimal health
B = time lived in suboptimal health
f(B) = function that assigns weights to years lived in suboptimal health (optimal health has a weight of 1)

Health gaps measure the difference between actual population health and some specified standard or goal (Equation 3). One of the most employed measures of health gaps is DALY measure (1,2).

$$HG = C + g(B) \qquad \textbf{Equation 3.}$$

HG = health gap
B = time lived in suboptimal health
C = time lost due to mortality (premature death)
f(B) = function that assigns weights to health states lived during time B, but where a weight of 1 equals to time lived in a health state equivalent to death.

Disability-Adjusted Life Year (DALY) concept

Introduction

DALY is an indicator of BoD in a population. It takes into account not only premature mortality, but also disability caused by disease or injury. As a new single summary measure was introduced in a 1990 Global Burden of Disease Study (GBDS) (6), which represented a major step in quantifying global and regional effects of diseases, injuries, and risk factors on population health. It is worth to note that DALYs are an inverse form of the more general concept of QALYs (7).

Two dimensions of DALYs

The DALY is a time-based measure that combines years of life lost due to premature mortality and years of life lost due to time lived in health states less than ideal health. One DALY can be thought of as one lost year of "healthy" life, and the BoD can be thought of as a measurement of the gap between current health status and an ideal situation where everyone lives into old age, free of disease and disability (1,4,8-11). In other words, DALYs are the combination (more precisely, the sum) of two dimensions: the present value of future years of lifetime lost through premature mortality, and the present value of years of future lifetime adjusted for the average severity (frequency and intensity) of any mental or physical disability caused by a disease or injury (10-12).

The years of life lost dimension

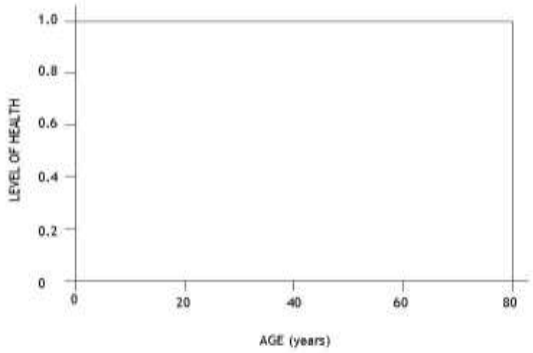
As a basis for the DALY measure, a "gold standard", or most desirable life, is defined as living in a completely healthy state until death at age around 80 years. Perfect health is 1 on the y-axis and death is 0 on the DALY diagram shown in Figure 2.

The "ideal" life is quantified as the total area in the box, a combination of the number of years lived and the full quality of life without disability (8). For each premature death[14], the number of years lost is counted up to the "standardised" maximum life span. The standardized maximum life span is 82.5 years for females and 80 years for males. It is taken from the country with the highest life expectancy in the world, Japan. Such a measure of premature

[14] Here, a premature death is defined as one that occurs before the age to which a person could have expected to survive assuming a life expectancy at birth approximately equal to that of the world's longest-surviving population - Japan (1,8,9).

death in number of years lost is known as "years of life lost" (YLL) (8). The following example illustrates how the YLLs could be calculated (Example 1).

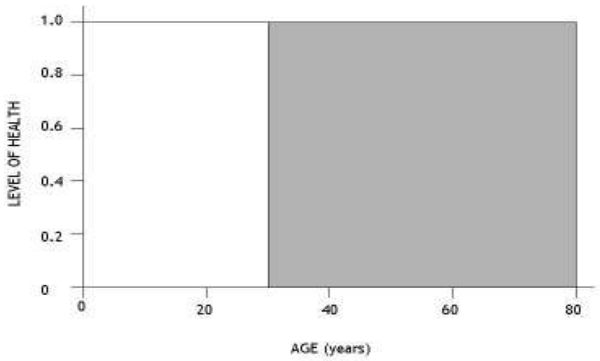
Figure 2. Graphical presentation of a life in full health until sudden death at the age of 80 years (8)



Example 1.

A scenario: a man dies in a car accident at 30 years of age.
In terms of years of life lost, 50 years are lost due to this premature death (YLLs = 80-30 years). This could be illustrated as presented in a Figure 3. The gray area represents the time lost due to premature death

Figure 3. Presentation of a life in full health until sudden death at the age of 30 years. The gray area represents the time lost due to premature death



The disability dimension

Injury and disease cause not only deaths but also varying time periods with morbidity and disability. The time period in years that is lived in states of poor health or disability due to each disease is another dimension of the DALY measure.

The disability is measured in length in years and in severity.

Severity weights have been appointed for each disabling condition on a scale from one to zero[15] (11,13,14). The disability severity weight for each disease reflects the average degree of disability a person suffers with each condition. Panels of healthy experts with knowledge about disease conditions have determined the weights. We will discuss the disability weights in details later in this module.

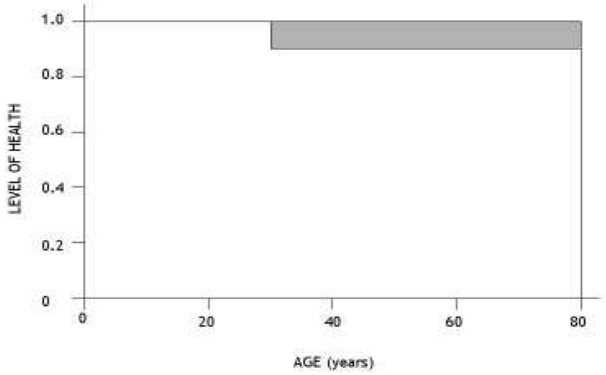
The severity weight is then multiplied by the average time a person is suffering from the disability from each disease (8). A measure of years lived in health states less than ideal health is known as "years lived with disability" (YLD) (8). Two examples from a person's life are presented in continuation (Examples 2 and 3).

Example 2.

A scenario: at the age of 30, a man gets a knee injury and his health is jeopardized with a weighted severity of 0.1. The injury is incurable and a man suffers until he dies at the age of 80 years.
In terms of years lost due to disability this man's health is only 0.9 of the maximum of 1.0 for the entire 50-year period. This could be illustrated (Figure 4). The grey area in Figure 4 represents his life years lost due to disability, and YLDs corresponds to 5 years (YLDs = 0.1×50 = 5 years).

[15] For example, schizophrenia was given a weighted severity loss of 0.8, whereas the common cold only causes a loss of 0.007.

Figure 4. Illustration of life of a man who gets a knee injury at the age of 30 years. The injury is incurable and a man suffers until he dies at the age of 80 years. The grey area represents life years lost due to disability

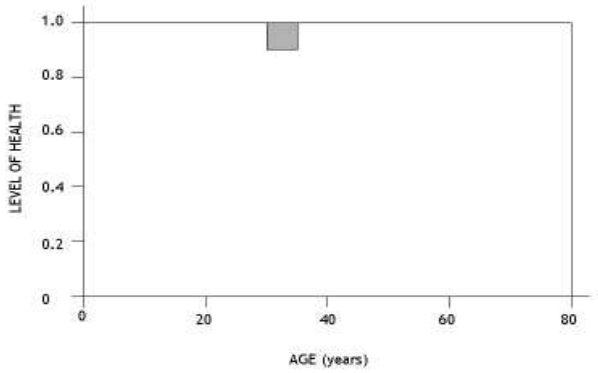


Example 3.

A scenario: at the age of 30, a man gets a knee injury and his health is jeopardized with a weighted severity of 0.1. During the years he suffers from the knee injury his health is only 0.9 of the maximum of 1.0. After At the age 35 he is successfully operated and recovers completely.

In terms of years lost due to disability this man suffers from the knee injury and his health is only 0.9 of the maximum of 1.0 for the 5-year period. This could be illustrated in a figure (Figure 5). The grey area in Figure 5 represents his life years lost due to disability, and YLDs correspond to 0.5 years (YLDs = 0.1×5 = 0.5 years).

Figure 5. Illustration of life of a man who gets a knee injury at the age of 30 years. At the age 35 he is successfully operated and recovers completely. A man is healthy until he dies at the age of 80 years. The grey area represents life years lost due to disability



Both dimensions combined

Usually, both dimensions are combined. Another two examples from a person's life are presented in continuation (Examples 4 and 5).

Example 4.

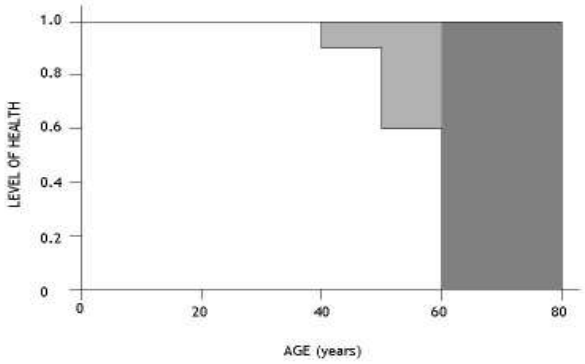
A scenario: at the age of 40, a man gets cancer which disables him for a certain amount but after a surgery he is in remission for 10 years. After 10 years he suffers from a progress of a disease which disables him substantially more. At the age of 60 years he dies.

In terms of years of life lost, 20 years are lost due to this premature death (YLLs = 80-60 years) (Figure 6, dark grey area).

In terms of years lost due to disability the health of this man is 0.9 for the 10-year period and after a progression is 0.6. YLDs in this case correspond to 5 years (YLDs = 0.1×10 + 0.4×10 = 1+4 = 5 years) (Figure 6, light grey area).

The total loss, that is DALYs, is 25 years (DALYs = YLDs+YLLs = 20+5 = 25)

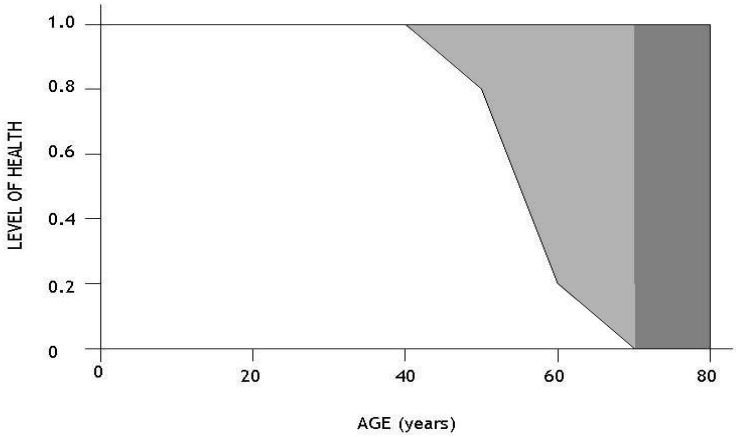
Figure 6. Illustration of a life of a man who gets a cancer at the age of 40 years. The disease disables him more and more. He dies at the age of 60 years. The grey area represents life years lost due to disability (light gray) and due to premature death (dark gray)



Example 5.

A scenario in this example is very similar as in Example 4. The difference is that the disease disables a man gradually more and more until he finally dies at the age of 70 years (Figure 7).

Figure 7. Illustration of a life of a man who gets a cancer at the age of 40 years. The disease disables him gradually more and more until he finally dies at the age of 70 years. The grey area represents life years lost due to disability (light gray) and due to premature death (dark gray)



The calculation in this case is more complex than presented in Examples 1-4.

Calculation of DALYs

The DALY measure is the sum of both dimensions/components just described - the sum of the YLLs and the YLDs (4,10,11,15-19) (Equation 4):

$$DALY = YLL + YLD \qquad \textbf{Equation 4.}$$

DALY= disability adjusted life years
YLL = years of life lost due to premature death
YLD = years lost due to disability

For YLLs we already know that they measure the number of years lost when a person dies prematurely. Consecutively, the younger is the age at which death occurs, the greater is the number of YLLs. The YLDs measure the number of years of healthy life lost due to temporary or permanent disability. Consecutively, the more severe disability or the longer the duration of this disability associated with a given health condition, the greater is the number of YLDs.

Another characteristic is that DALYs basically comprise so called social values/preferences. The two basic ones are:

- the sex differences which are built in the YLL component (standard life expectancies are separate for men and for women), and
- disability weighting which is built in the YLDs component.

But they could comprise additional social preferences, for example discounting and age weighting as well. According to this characteristic, DALYs could be computed:

- without considering age weighting and discounting,

- with considering age weighting, or
- with considering age weighting and discounting.

All procedures will be discussed in continuation but only the most simple, without considering age weighting and discounting will be presented in details since both other procedures are out of the scope of this module.

Basic procedure considering basic social preferences

Before presenting the basic procedure of calculation of DALYs the main characteristics of two basic social preferences built in the DALY calculation procedure need to be discussed in more details.

Two basic social preferences - sex differences and disability weighting

Major characteristics of basic social preferences are:

1. **Sex differences**

To assess premature mortality, a standard life table for all populations, with life expectancies at birth fixed at 82.5 years for women and 80 years for men is utilized. Life expectancy is calculated separate for men and for women because women live on average longer than men. This table could be accessed in full or abridged form from the publicly available World Health Organization (WHO) web page[16] (20,21). It is also presented in Appendix (Appendix, Table A1). The same table was used by Murray and co-workers (22). What this table tells us is illustrated in Example 6.

Example 6.

Standard life expectancy table tells that (23):

- a male infant who dies shortly after birth would lose all 80.00 years of his life he would have been expected to live (Appendix, Table A1),
- a man who dies at age 50 would lose 30.99 years of expected life, because if he has already survived up to age 50 he has a life expectancy of 80.99 years (50 years + 30.99 years) (Appendix, Table A1),
- a man who dies at age 75 would lose only 10.17 years of expected life, because his life expectancy at age 75 is 85.17 years (75 years + 10.17 years) (Appendix, Table A1).

A standard life expectancy allows deaths at the same age to contribute equally to the BoD irrespective of where the death occurs. Alternatives, such as using different life expectancies for different populations that more closely match their actual life expectancies, violate this egalitarian principle. As life expectancy is rarely equal for men and women, a lower reference life expectancy could be used for both sexes because of "biological differences in survival potential", as well as men are higher exposed to various risks such as alcohol, tobacco, and occupational injury (1,9,24).

2. **Disability weights**

On one side, there are health conditions that frequently cause significant disability or death, while on the other side there are those conditions that rarely cause death but may cause severe and/or prolonged disability. The other health conditions can cause severe disability but they occur late in life and they are of shorter duration (Example 7).

Example 7.

Here we give some examples:

- *Alzheimer's disease often ca*use severe disability but it typically occurs in older ages and consecutively lasts shorter,
- Down syndrome is a congenital anomaly that causes limited ability to perform most activities for entire life of a person,
- stroke is highly lethal disease etc.

In order to quantify time lived with a non-fatal health outcome and assess disabilities in a way that will help to inform health policy, disability must be defined, measured, and valued in a clear framework that inevitably involves simplifying reality.

Disabilities were assigned severity weights ranging from zero, representing perfect health, to one, representing death. These weights were determined at a meeting of experts in international health who had not participated in the first GBDS. In order to reduce the number of weights to be assigned and to emphasize large differences in the severity of disability, each disability condition was assigned to one of six classes. In Table 1 definitions of each of these classes are presented (22).

[16] The GBD methods can be applied on a national or sub-national level and tailored to health values of that country. WHO provides practical tools for people undertaking a BoD study. A link to software to help create life tables and disease models, a manual, supplementary files to help with calculations, and examples of where these methods have been applied by researchers around the globe.

In continuation, for disability, the incidence of cases by age, sex, and demographic region was estimated on the basis of community surveys or, failing that, on expert opinion; the number of years of healthy life lost was then obtained by multiplying the expected duration of the condition (up to remission or to death) by a severity weight that measured the severity of the disability in comparison with loss of life.

For the GBDS 2000 and 2004 some updates were made and diseases were grouped into seven classes of severity of disability. For example, class IV, which includes arm or below-the-knee amputation, deafness etc., was given a range of severity weight from 0.24-0.36, and class VI, which includes AIDS cases not on antiretroviral drugs, Alzheimer and other dementias, blindness etc., was assigned a severity weight of 0.6 (range from 0.5-0.7) (Table 2) (11,14).

Two methods are commonly used to formalize social preferences for different states of health. Both involve asking people to make judgements about the trade-off between quantity and quality of life. This can be expressed as a trade-off in time (how many years lived with a given disability a person would trade for a fixed period of perfect health) or a trade-off between persons (whether the person would prefer to save 1 year of life for 1,000 perfectly healthy individuals or 1 year of life for perhaps 2,000 individuals in a worse health state). The DALY approach which is the basis of the GBDS currently in use has been much criticised because the method presupposes that life years of disabled people are worth less than life years of people without disabilities (1,9). The GBDS for example developed a protocol based on the person trade-off method. In a formal exercise involving health workers from all regions of the world, the severity of a set of 22-indicator disabling conditions - such as blindness, depression, and conditions that cause pain - was weighted between 0 (perfect health) and 1 (equivalent to death). These weights were then grouped into seven classes where class 1 has a weight between 0 and 0.02 and class VII a weight between 0.70 and 1 (Table 2). In the protocol, a life year for 1,000 healthy people has been set as equally valuable as one life year for (1,5):

- 9,524 people with quadriplegia;
- 2,660 blind people;
- 1,686 people with Down's syndrome without cardiac malformation;
- 1,499 deaf people;
- 1,236 infertile people;

WHO has announced changes to such an approach.

Classification for estimating mortality and disability

Deaths were classified using a tree structure, in which the first level of disaggregation comprises three broad cause categories of diseases (1,9,14,16):

- Group I: communicable diseases, perinatal, and nutritional conditions;
- Group II: non-communicable diseases;
- Group III: injuries.

Each group was then subdivided into categories: for example, cardiovascular diseases and malignant neoplasms are two subcategories of group II. Beyond this level, there are two further disaggregation levels such that 107 individual causes from the ninth revision of the ICD (ICD-9) can be listed separately. Consistent with the goal of providing disaggregated estimates of BoD to assist priority setting in the health sector, estimates were prepared by age and sex and for eight broad geographic regions of the world: Established Market Economies, Formerly Socialist Economies of Europe, China, India, Latin America and the Caribbean, Middle-Eastern Crescent, Other Asia and Islands, and sub-Saharan Africa (1,9).

The disabilities in a particular class differ in kind (for example, blindness versus paralysis) but were considered to be of equal severity. Each participant then voted on the weight to be assigned to the entire class, not to individual disabilities, and the class was weighted according to the average vote. It is important to note that many disabling conditions lead to two or more distinct disabilities, which may be classified in more than one class of severity (9,11,24).

There is surprisingly wide agreement between cultures on what constitutes a severe or a mild disability. For example, a year lived with blindness appears to most people to be a more severe disability than a year lived with watery diarrhea, while quadriplegia is regarded as more severe than blindness. These judgements must be made formal and explicit if they are to be incorporated into measurements of BoD (1,9,14).

When one would like to perform calculation of YLD, he/she will need to find the actual list of disability weights. In the GBDS 2004 update publication is stated that the authors used for the weight factor the weights listed in Annex Table A6 of a publication of Mathers et al (7). The same table could be obtained as special WHO document available on the Internet (25), while more detailed table, including weights according to age, is available on special WHO web page on the Internet (20,26). For the purposes of this module an adapted table presenting average disability weights for diseases and conditions including cancers and injuries is available in Appendix (Appendix, Table A2).

Table 1. Definitions of disability weighting classes. Adapted from Murray (22)

Class	Description	Weight
1	Limited ability to perform at least one activity in one of the following areas: recreation, education, procreation or occupation	0.096
2	Limited ability to perform most activities in one of the following areas: recreation, education, procreation or occupation	0.220
3	Limited ability to perform activities in two or more of the following areas: recreation, education, procreation or occupation	0.400
4	Limited ability to perform most activities in all of the following areas: recreation, education, procreation or occupation	0.600
5	Needs assistance with instrumental activities of daily living such as meal preparation, shopping or housework	0.810
6	Needs assistance with activities of daily living such as eating, personal hygiene or toilet use	0.920

Table 2. Disability classes for the Global Burden of Disease Study (GBDS), with examples of long-term disease and injury sequelae falling in each class[1] (11,14)

Disability class	Severity weights	Conditions[2]
I	0.00-0.02	Stunting due to malnutrition, schistosomiasis infection, long-term scarring due to burns (less than 20% of body)
II	0.02-0.12	Amputated finger, asthma case, edentulism, mastectomy, severe anaemia, stress incontinence, watery diarrhea
III	0.12-0.24	Angina, HIV not progressed to AIDS, alcohol dependence and problem use, radius fracture in a stiff cast, infertility, erectile disfunction, rheumatoid arthritis, angina, low vision (<6/18, >3/60)
IV	0.24-0.36	Amputated arm, below-the-knee amputation, deafness, congestive heart failure, drug dependence, Parkinson disease, tuberculosis
V	0.36-0.50	Bipolar affective disorder, rectovaginal fistula, mild mental retardation, neurological sequelae of malaria
VI	0.50-0.70	AIDS cases not on antiretroviral drugs, Alzheimer and other dementias, blindness, paraplegia, Down syndrome,
VII	0.70-1.00	Active psychosis (schizophrenia), severe depression, severe migraine, quadriplegia, terminal stage cancer

[1] Based on average severity weight globally for both sexes and all ages in the GBDS 2004 update.
[2] Conditions are listed in the disability class for their global average weight. Most conditions will have distributions of severity spanning more than one disability class, potentially up to all seven.

The basic procedure

We will present the basic procedure for calculating DALYs by using an example (Example 8).

Example 8. A scenario: a woman who had moderate depression since she was 20 commits a suicide at age 50.

1. Calculating the YLLs

On a population basis the YLLs for a given age basically correspond to the number of deaths for that given age multiplied by the standard life expectancy at the age at which death occurs. The basic formula for calculation of YLLs on a population basis is the following (12,15,19) (Equation 5):

$$YLL = N \times L \qquad \textbf{Equation 5.}$$

YLL = years of life lost due to premature death
N = number of deaths
L = standard life expectancy at age of death in years

On an individual basis the YLLs for an individual person correspond to the standard life expectancy at the age at which death occurs. We have already worked out this situation in Example 6, but it is worked out again for Example 8. For the woman from Example 8 the YLLs are calculated as follows:

In terms of YLLs the woman from scenario presented in Example 8: **Example 8. Cont.**

- would lose 33.99 years of expected life, because if she has already survived up to age 50 she has a life expectancy of 83.99 years (50 years + 33.99 years) (Appendix, Table A1),
- YLL = 33.99

2. Calculating the YLDs

Because YLLs measure the incident stream of lost years of life due to deaths, an incidence perspective is also taken for the calculation of YLDs. To estimate YLDs for a particular cause in a particular time period, the number of incident cases in that period is multiplied by the average duration of the disease and a weight factor that reflects the severity of the disease on a scale from 0 (perfect health) to 1 (death). The basic formula for calculation of YLDs on a population basis is the following (12,15,19) (Equation 6):

$$YLD = I \times DW \times L \qquad \textbf{Equation 6.}$$

YLD = years lost due to disability
I = number of incident cases
DW= disability weight
L = average duration of the case until remission or death (years)

On an individual basis the basic formula for calculation of YLDs is the following (Equation 7):

$$YLD = DW \times L \qquad \textbf{Equation 7.}$$

YLD = years lost due to disability
DW= disability weight
L = duration of the case until remission or death (years)

For the woman from Example 8 the YLDs are calculated as follows:

In terms of YLDs: **Example 8. Cont.**

- her disability is weighted to 0.350 (disability weight for moderate depressive episode) (Appendix, Table A2),
- it lasts for 30 years (50-20 years),
- YLD = (20×0) + (30×0.350) = 10.50 years

3. Calculating DALYs

At the end the YLLs and the YLDs are summed up according to Equation 4. For the woman from Example 8 the DALYs are calculated as follows:

DALY = 33.99 + 10.50 = 44.49 **Example 8. Cont.**

The burden of disease in this case in terms of DALYs is 44.49 years.

Procedures considering additional social preferences

Before presenting procedures of calculation of DALYs that consider additional social preferences their main characteristics are briefly discussed.

Additional basic social preferences – discounting and age weighting

Major characteristics of additional social preferences are:

1. Discounting

Discounting means that future gains and losses are counted less than if they had occurred today. The years lost in the future are discounted, so that years lost now are worth more than years lost in the future. This is a standard procedure and common practice in economics when it comes to valuing material goods, and in the DALY calculations a discount rate of 3% per year is used (8,9,16,27).

The innovative GBDS for example calculated the total sum of the combined loss of all premature deaths that occurred in the world in 1990 and the loss of healthy life from disability in future years from specific diseases arising in that year (28). The study used all possible data sources of recorded causes of death and prevalence and incidence of disease, as well as expert judgment when data were not available (8).

Individuals commonly discount future benefits against current benefits similarly to the way that they may discount future dollars against current dollars. Whether a year of healthy life, like a dollar, is also deemed to be preferable now rather than later, is a matter of debate among economists, medical ethicists, and public health planners, since discounting future health affects both measurements of disease burden and estimates of the cost-effectiveness of an intervention (3,12). There are arguments for and against discounting. In the GBDS, for example, future life years were discounted by 3% per year. This means that a year of healthy life bought for 10 years hence is worth around 24% less than one bought for now, as discounting is represented as an exponential decay function. Another effect is that it reduces the value of interventions, especially

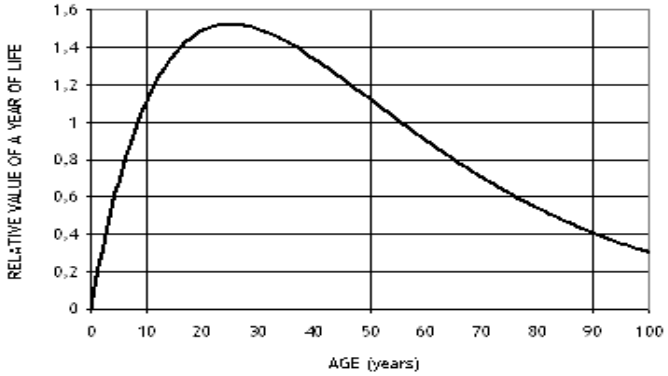
preventive, that provide benefits largely in the future, such as vaccinating against hepatitis B, which may prevent thousands of cases of liver cancer, but some decades later (1,9,11,12).

2. **Age weighting**

Another characteristic of the DALY measure is that years lost due to premature deaths or disability could give different values at different ages. These differences in values are introduced in DALY calculations by what is called "age weights"'. The age weight used in the DALY calculations is obtained from a scale where the value of a year lost rises steeply from zero at birth to a maximum at 25 years of age, and then decreases progressively in older ages (Figure 8). From the Figure 8 it could be seen that the relative value of a life year is below 1 for the children under 10, and for the persons more than about 55 years of age.

Because of weighting the value of the lifetime the years of life in childhood and old age are counted less because of social roles vary with age and social value of that time may be different i.e. "young, and often elderly, depend on the rest of society for physical, emotional and financial support" (3,24). This means that if a newborn girl dies, 32.5 weighted years are lost, if she dies at age 30, 29 weighted years are lost, and at age 60, 12 weighted years are lost (8). For males, the above figures will be slightly lower, because their shorter life expectancy is taken into consideration. Another example is that living with disability, e.g. blindness, for a person aged 80 years is considered "less bad" than living with blindness for a 25-year old individual (1,9,16,24).

Figure 8. Relative values of a year of life. Adapted from the World Bank development report 1993 (9)



If individuals are forced to choose between saving a year of life for a 2-year old and saving it for a 22-year old, most prefer to save the 22-year old one. A range of studies confirms this broad social preference to weight the value of a year lived by a young adult more heavily than one lived by a very young child, or an older adult. Adults are widely perceived to play a critical role in the family, community, and society. It was for these reasons that the GBDS incorporated age weighting into the DALYs. It was assumed that the relative value of a year of life rises rapidly from birth to a peak in the early twenties, after which it declines steadily (1,3,8).

Calculation of DALYs with discounting

When 3% discounting and uniform age weights are considered, the formulas for calculating YLLs (Equation 8) and YLDs (Equation 9) on the population basis are the following (19,29):

$$YLL = \frac{N}{0.03}\left(1 - e^{-0.03L}\right) \qquad \textbf{Equation 8.}$$

YLL = years of life lost due to premature death
N = number of deaths
L = standard life expectancy at age of death in years

$$YLD = \frac{IDW\left(1 - e^{-0.03L}\right)}{0.03} \qquad \textbf{Equation 9.}$$

YLD = years lost due to disability
I = number of incident cases
DW= disability weight
L = average duration of the case until remission or death (years)

Detailed presentation of this procedure is out of the scope of this module.

Calculation of DALYs with discounting and age weighting

When non-zero discounting and age weighing are considered, the formulas for calculating YLLs (Equation 10) and YLDs (Equation 11) on an individual basis are the following (12,19,29):

$$YLL = \frac{KCe^{ra}}{(\beta + r)^2}\begin{bmatrix}\left[e^{-(r+\beta)(L+a)}\right]\\ \left[-(r+\beta)(L+a)-1\right]\\ \left[-e^{-(r+\beta)a}\right]\\ \left[-(r+\beta)a-1\right]\end{bmatrix} + \left[\frac{1-K}{r}\left(1-e^{-rL}\right)\right] \qquad \textbf{Equation 10}$$

YLL = years of life lost due to premature death
K = age-weighting modulation constant (e.g. K= 1)
C = adjustment constant for age weights (GBDS standard value is 0.1658)
r = discount rate (GBDS standard value is 0.03)
a = age of death (years)
ß= age-weighting constant (GBDS standard value is 0.04)
L = standard life expectancy at age of death (years)

$$YLL = DW\left\{\frac{KCe^{ra}}{(\beta + r)^2}\begin{bmatrix}\left[e^{-(r+\beta)(L+a)}\right]\\ \left[-(r+\beta)(L+a)-1\right]\\ \left[-e^{-(r+\beta)a}\right]\\ \left[-(r+\beta)a-1\right]\end{bmatrix} + \left[\frac{1-K}{r}\left(1-e^{-rL}\right)\right]\right\} \qquad \textbf{Equation 11.}$$

YLD = years lost due to disability
DW= disability weight
K = age-weighting modulation constant (e.g. K= 1)
C = age-weighting correction constant (GBDS standard value is 0.1658)
r = discount rate (GBDS standard value is 0.03)
a = age of onset
ß = parameter from the age-weighting function (GBDS standard value is 0.04)
L = duration of disability (years)

Detailed presentation of this procedure is out of the scope of this module since it is very complicated and a separate module is needed to present it. For all who want to study this procedure in more details, a paper by Fox-Rushby and Hanson (12), is recommended where also the formulas in Excel programme for calculation of YLLs and YLDs are given.

These formulas have also been programmed into calculation spreadsheet templates for calculation of DALYs on a population basis (19,29), available at the WHO web site (20,30).

Sensitivity analysis

To gauge the impact of changing these social choices on the final measures of BoD, the GBDS assessments were recalculated with alternative age weighting and discount rates, and with alternative methods for weighting the severity of disabilities. Overall, the rankings of diseases and the distribution of burden by broad cause groups are largely unaffected by age weighting and only slightly affected by changing the method for weighting disability. Changes to the discount rate, by contrast, may have a more significant effect on the overall results. Changes in the age distribution of burden, in turn, affect the distribution by cause, as communicable and perinatal conditions are most common in children while non-communicable diseases are most common in adults. The most significant effect of changing the discount and age weights is a reduction in the importance of several psychiatric conditions (1,11,24).

However, sensitivity analysis has shown that the results of the GBDS are not greatly affected by these social preferences. Another problem is that the GBDS calculates DALY on data, which on some continents are of poor quality. Especially for the disability calculations, the data is of varied quality in different regions, e.g. Sub-Saharan Africa, and for different disease conditions, e.g. depression (8,11).

Controversies and criticism

Some critical articles on the DALY approach have questioned both the validity of the results as well as the underlying value-judgements (24,27). In the Journal of Health Economics Anand and Hanson argues that: "the conceptual and technical basis for DALYs is flawed, and that the assumptions and value judgements underlying it are open to serious question" (24).

According to some authors, the DALY concept has "the potential to revolutionize the way in which we measure the impact of disease, how we choose interventions, and how we track the success or failure of our intervention (31,32). Furthermore, DALYs are considered to be an "advancement" over other composite indicators, such as QALYs, because the value choices incorporated in the DALYs are made explicit: "The black box of the decision-maker's relative values is then opened for public scrutiny and influence" (24). Yassin (33), pointed out several advantages for using DALY in studies of health inequalities:

- the DALY is the only measure that can infuse information about non-fatal health outcomes into debates of health inequalities;
- DALY uncouples social and epidemiological assessment of health inequalities from advocacy;
- the DALY can measure the magnitude of premature death and non-fatal health outcomes attributable to proximal biological causes, including diseases and injuries or attributable to more distal causes such as poor living standards, tobacco use or socio-economic determinants;
- the DALY is a stable measure that can be used for purposes of comparisons either between different communities or between different points of time.

Jankovic (3) emphasized that DALY measurement of clinical outcomes and cost-effectiveness analyses allows existing or prospective interventions to be judged both in terms of cost-effectiveness, and their relative impact in reducing the BoD and ill-health. DALY as a composite indicator is a useful analytical tool for health policy-makers and analysts in priority setting and resource allocation in health systems providing unique and desirable health information on non-fatal health outcomes that is essential for determining appropriate health research priorities, too (3).

The DALY measure has been criticized because of the four built-in social preferences:

- different weights for sexes;
- different age weights;
- discounting future years lost, and;
- severity weighting of disabilities.

Many argue that life years for men and women should be given the same weight. However, as has been described above the difference is small and only gives a slightly greater value for diseases that affect females. Some people argue that all years lost should be given the same value independently of the age at which the years are lost. Others argue that discounting years is wrong, because they value the current and the future years equally. In a complex measure like DALY, the built-in social preferences may conceal issues of inequity. The most difficult part of any approach combining data on quality of life and length of life is how to measure the quality of life. Many philosophical questions as well as questions regarding the limits of natural sciences arise. The first requirement of a valid measurement is that one knows what is being measured. The concept of quality of life is, however, vaguely defined, and different people as well as different cultures may have very different opinions on the main elements of a good life (16).

The approach has been criticized for violating the principle of treating people equally and for discriminating the young, the elderly, future generations (future health benefits), the disabled, and the women (16,24).

Use of DALYs

This approach increased the validity of comparisons of the burden of different diseases between world regions and countries over time. In fact, the World Bank and the World Health Organization (WHO) were the first institutions to use the DALY measures to compare the BoD in different regions of the world and thereby the value and effectiveness of different health interventions and changes in living conditions. It became possible to estimate and compare the cost of avoiding the loss of a DALY for each intervention (4,8,9).

Prior to the GBDS, which began in 1992, there had been no comprehensive efforts to provide comparable regional and global estimates and projections of the causes of loss of health and disease and injury burden in populations based on a common methodology and denominated in a common metric comparable across populations and over time (1,9).

One of the major goals of the GBDS was to facilitate the inclusion of non-fatal illness/conditions and their long-term health consequences/outcomes (mental and musculoskeletal disorders, blindness etc.) into debates on international health policy, beside the causes of death and mortality data. In addition, there was a need to quantify the BoD using a measure that could then be used for cost-effectiveness analysis. The GBDS method quantifies not merely the number of deaths but also the impact of premature death and disability in a population, combining these measures into a single unit of measurement of the overall BoD in the population - the DALY. DALYs allow the losses or the BoD from the premature death and nonfatal consequences of over 100 diseases and injuries to be expressed in the same unit. The study also presented the first global and regional estimates of disease and injury burden attributable to certain risk factors for disease, such as tobacco, alcohol, poor water and sanitation, and unsafe sex. The method uses 107 diagnoses, covering all conceivable causes of death and 95% of all possible causes of disability (1).

The methods of the GBDS 1990 created a common metric system to estimate the health loss associated with morbidity and mortality. It generated widely published findings and comparable information on disease and injury incidence and prevalence for all world regions. It also stimulated numerous national studies of BoD. These results have been used by governments and non-governmental agencies to inform priorities for research, development, policies and funding. In 2000, the WHO began publishing regular GBDS updates for the world and 14 regions. These revisions were aided by methodological improvements and more extensive data collection that covered key aspects of the global BoD, including mortality estimation, cause of death analysis, and measurement and valuation of

functional health status. Standardized concepts and approaches to comparative risk assessment were applied for over 25 risk factors (10,11).

Conclusion

Bearing its inherent weaknesses in mind, DALY is still a very useful measure, because it is the first comprehensive attempt made to summarise the world's burden of injury, disease and premature death. It has initiated a debate and new research to find even better complex indicators for global comparisons (34). The DALY measure is useful to describe the disease burden across the world and to make projections for the future. At present, many countries are exploring the possibility of using DALYs as a measure of trends in disease burden and as a tool for cost-effectiveness studies and priority setting (8,32).

In 1999 WHO has started to include the DALY measure in their annual reports. This allows for refinement of the results from the initial study because of new health data. From the year 2000, they also included a DALE, disability-adjusted life expectancy, which was renamed as the more cheery HALE, health-adjusted life expectancy, in 2002. This measure is based on life expectancy at birth, but includes an adjustment for the time spent in poor health. It is the equivalent of the number of years a newborn can expect to live in full health, based on current statistics of mortality and morbidity. In Japan, for instance, the HALE is 72 years, while in Afghanistan only 35 years. Many find this a measure that is instinctively easier to understand compared to DALY (3,8).

Case study

A scenario

In a heavy frontal car collision of two cars 7 people are involved. In the first car a 4-member family dies, while in the second car 3 young people are heavily injured. The sequelae of a car collision for every participant are presented in Table 3.

Table 3. The sequelae of a car collision for every participant with their weights

	Participant	Sequelae	Disability weight
	CAR 1		
1.	36 years of age father	Dies	1
2.	29 years of age mother	Dies	1
3.	7 years of age daughter	Dies	1
4.	2 years of age son	Dies	1
	CAR 2		
5.	27 years of age man	Injured spinal cord (lifelong)	0.725
6.	25 years of age woman	Intracranial injury - short term (½ year)	0.359
7.	22 years of age woman	Fractured ribs (¼ year)	0.199

Calculation (basic procedure)

In continuation, we will calculate the DALYs for this car accident. In Table 4 is presented calculation of elements of DALYs - YLLs and YLDs - and the final result (DALYs).

Table 4. Elements of calculation of DALYs and final result

Participant	YLLs	YLDs	DALYs
1.	44.58 years LE		
2.	54.25 years LE		
3.	75.97 years LE		
4.	78.36 years LE		
5.		53.49 years LE × 0.725 = 38.78	
6.		0.5 × 0.359 =0.18	
7.		0.25× 0.199 =0.05	
Total	253.16	39.01	292.17

LEGEND: DALY = disability adjusted life years; YLL = years of life lost due to premature death; YLD = years lost due to disability; LE = life expectancy.

In summary, altogether 292.17 DALYs are lost in this heavy frontal car collision.

Exercises

Task 1

After the introductory lecture, students carefully read the part on theoretical background of this module and corresponding recommended readings.

Task 2

Students consider the following scenario: at the age of 55 years, a previously healthy man is diagnosed with prostate cancer. After a surgery and radiotherapy he is in remission for 15 years. At the age of 70 years, he is diagnosed with a metastatic disease. He dies at the age of 75 years.

In groups of 2-3, in the process of calculation of DALYs, students should conduct the following steps:

- make graphical presentation (a sketch) of a time horizon for this case;
- calculate DALYs[17];
- compare own results with the results of other student groups.

Task 3

This task is based on the examples presented in the manual "Selecting an essential package of health services using cost-effectiveness analysis: a manual for professionals in developing countries" (35).

Let us imagine a 5-year old girl who falls sick with poliomyelitis at this age. Following scenarios are possible:

1. immediately after she contracts poliomyelitis at age 5, she dies;
2. after she contracts the disease at age 5 she lives until age 10 with a disability;
3. after she contracts the disease at age 5 she becomes permanently disabled over her entire life span (77.95 years);
4. after she contracts the disease at age 5 she is in an acute phase of the disease confined to bed for ½ month (disability weight 0.500). Afterwards she fully recovers.

In the process of calculation of DALYs, students should conduct the following steps (this task is to be completed by every student alone):

- make graphical presentation (a sketch) of a time horizon for each of these scenarios;
- calculate DALYs (be careful in calculating YLLs and YLDs; help yourself with graphical presentations) for each scenario[18];
- compare own results with the results of other students.

Task 4

This task is continuation of the Task 3 and is based on working paper of Homedes (36).

Let us imagine that in a particular community there are 20 girls who fall sick with poliomyelitis at age 5:

- 2 of them die immediately;
- 2 die at age 10 after a 5-year period of disability;
- 10 of them are permanently disabled, and;
- 6 recover completely after the ½ month acute phase of the disease.

Calculate the total number of DALYs lost due to poliomyelitis in this community[19].

Task 5

Students first carefully study the data presented in Table 5.

Afterwards they should consider following tips:

- they discuss what these data present;
- they calculate BoD (expressed as DALYs) using the data from the Table 5;
- they should use the basic formula for a simple way of calculation, without age-weighting and discounting;
- they should compare the results[20];
- if possible, students should calculate DALYs for various diseases and injuries causing disability and/or premature death, using official data from their own countries, or local communities.

[17] Final result: 31.93 years

[18] Final results:
Scenario 1: 77.95 years
Scenario 1: 74.835 years
Scenario 1: 28.764 years
Scenario 1: 0.021 years

[19] Final result: 593.336 years

[20] Final result: 2,808 years

Table 5. Morbidity and mortality from ishaemic heart disease (heart attack), by the age of getting heart attack and period of survival, in the population of the Region X in SEE, 2008

Health outcome of the heart attack		Age group							
		< 35	35-39	40-44	45-49	50-54	55-59	60-64	65 +
Deaths	No. of cases	9	14	16	13	12	10	8	7
	YLL per case	40	35	30	25	20	15	10	5
Disability (weight 0.3)	No. of cases	9	14	16	13	12	10	8	7
	Years of survival	40	35	30	25	20	15	10	5

Task 6

This task comprises the following steps:

- in PubMed Central, students should try to find papers using DALYs as summary measures of BoD;
- students should analyze the ratio of DALYs calculated according to the basic method: DALYs calculated using discounting or/and age weighting;
- students should critically discuss the pros and cons of these different methods.

References

1. Murray CJL, Lopez AD. Assessing health needs: the Global Burden of Disease Study. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford Textbook of Public Health - Fourth edition. New York: Oxford University Press Inc.; 2004. p.243-54.
2. Atanackovic-Markovic Z, Bjegovic V, Jankovic S, et al. The burden of disease and injury in Serbia. Serbian Burden of Disease study - an EU funded project managed by the European Agency for Reconstruction. Belgrade: Ministry of Health of the Republic of Serbia; 2003.
3. Jankovic S. Summary measures of population health and their relevance for health policy. In: Galan A, Scintee G, editors. Public Health Strategies. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2005. p.190-207. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: June 5, 2013.
4. Tulchinsky TH, Varavikova EA. Measuring and evaluating the health of a population. In: Tulchinsky TH, Varavikova EA. The New Public Health. An Introduction for the 21st century. San Diego: Academic Press; 2000. p.113-68.
5. Arnesen TM, Norheim OF. Disability Adjusted Life Years - possibilities and problems. National Institute of Public Health, Oslo, Norway. Available from URL: http://iier.isciii.es/supercourse/lecture/lec2911/index.htm. Accessed: June 5, 2013.
6. World Health Organization. About the Global Burden of Disease (GBD) project [homepage on the Internet]. Available from URL: http://www.who.int/healthinfo/global_burden_disease/about/en/index.html. Accessed: June 5, 2013.
7. Mathers CD, Lopez AD, Murray CJL. The burden of disease and mortality by condition: data, methods and results for 2001. In: Lopez AD, Mathers CD, Ezzati M, Murray CJL, Jamison DT, eds. Global burden of disease and risk factors. New York, Oxford University Press; 2006. p.45-240.
8. Lindstrand A, Bergstrom S, Rosling H, Rubenson B, Stenson B, Tylleskar T. Global health - an introductory textbook. Copenhagen: Narayana Press; 2008.
9. The World Bank. The World Bank development report 1993. Investing in Health. New York, NY: Oxford University Press; 1993.
10. Institute for Health Metrics and Evaluation. The Global Burden of Disease: Generating Evidence, Guiding Policy. Seattle, WA: IHME, 2013.
11. World Health Organization. Global Burden of Disease 2004 Update. Geneva: World Health Organization; 2004.
12. Fox-Rushby JA, Hanson K. Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis. Health Policy Plann. 2001;16:326-31.
13. Stibich M. Disability Adjusted Life Years - DALYs - Measuring Suffering. Available from URL: http://longevity.about.com/od/longevity101/g/dalys.htm Accessed: June 5, 2013.

14. Galan A. Disability-Adjusted Life Years: A Method for the analysis of the burden of disease. In: Bjegovic V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.140-50. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: June 5, 2013.
15. World Health Organization. Metrics: Disability-Adjusted Life Year (DALY) - quantifying the burden of disease from mortality and morbidity. Available from URL: http://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/index.html. Accessed: June 5, 2013.
16. Holden B. DALYs and disagreement. Available from URL: http://blog.givewell.net/?p=269 Accessed: June 5, 2013.
17. Somerford P, Katzenellenbogen J, Western Australian Burden of Disease Study: Disability-Adjusted Life Years: Technical overview. Perth: Department of Health; 2004. Available from URL: http://www.health.wa.gov.au/publications/documents/BOD/BOD9.pdf. Accessed: June 5, 2013.
18. World Health Organization. Global Burden of Disease (GBD). Available from URL: http://www.who.int/healthinfo/global_burden_disease/en/index.html. Accessed: June 5, 2013.
19. Pruss-Ustun A, Mathers C, Corvalan C, Woodward A. Introduction and methods: assessing the environmental burden of disease at national and local levels. WHO Environmental Burden of Disease Series, No. 1. Geneva: World Health Organization; 2003.
20. World Health Organization. National tools [homepage on the Internet]. Available from URL: http://www.who.int/healthinfo/global_burden_disease/tools_national/en/. Accessed: June 5, 2013.
21. World Health Organization. National tools. National burden of disease supplementary files. Standard life table. Available from URL: http://www.who.int/healthinfo/bodreferencestandardlifetable.xls. Accessed: June 5, 2013.
22. Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ. 1994;72:429-445.
23. Los Angeles County Department of Health Services, UCLA Center for Health Policy Research. The burden of disease in Los Angeles County. Available from URL: www.lapublichealth.org/epi/reports/dburden.pdf. Accessed: June 5, 2013.
24. Anand S, Hanson K. Disability adjusted life years: a critical review. Journal of Health Economics, 1997;16:658-702.
25. World Health Organization. Global burden of disease 2004 update: disability weights for diseases and conditions. Geneva: World Health Organization; 2004.
26. World Health Organization. National tools [homepage on the Internet]. Available from URL: http://www.who.int/healthinfo/bodreferencedisabilityweights.xls. Accessed: June 5, 2013.
27. Murray CJL, Salomon JA, Mathers CD, Lopez AD (Eds.). Summary measures of population health. concepts, ethics, measurement and applications. Geneva: World Health Organization, 2002.
28. Murray CJL, Lopez AD (Eds.). The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2030. Global Burden of Disease and Injury Series, Vol. 1. Cambridge, MA: Harvard University Press; 1996.
29. World Health Organization, Global Program on Evidence for Health Policy. National burden of disease studies: a practical guide. Geneva: World Health Organization; 2001.
30. World Health Organization. DALY calculation template. Available from URL: http://www.who.int/healthinfo/bodreferencedalycalcSulationtemplate.xls. Accessed: June 5, 2013.
31. Foege W. Preventive medicine and public health. Journal of American Medical Association, 1994;271:1704-5.
32. Bobadilla JL, Cowley P, Musgrove P, Saxenian H. Design, Content and Financing of an Essential National Package of Health Services. The World Bank Group Private and Public Initiatives: Working Together in Health and Education. The World Bank, March 17,1995:1-10.
33. Yassin K, Galan A. Inequalities in Health as Assessed by the Burden of Disease Method - Case Study. In: Bjegovic V, Donev D, eds. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company, 2004:163-9.
34. Murray CJ, Lopez AD. Progress and directions in refining the global burden of disease approach: a response to Williams. Health Economics. 2000;9:69-82.
35. Brenzel L. Selecting an essential package of health services using cost-effectiveness analysis: a manual for professionals in developing countries. Washington, DC, Boston, MA: The World Bank, Population,

Health and Nutrition Department; Harvard School of Public Health, Department of Population Studies and International Health; 1993.

36. Homedes N. The disability-adjusted life year (DALY): definition, measurement and potential use. Human Capital Development and Operations Policy Working Paper. Washington DC: The World Bank; 1995.

Recommended readings

1. Jankovic S. Summary measures of population health and their relevance for health policy. In: Galan A, Scintee G, editors. Public Health Strategies. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2005. p.190-207. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: June 5, 2013.
2. Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ. 1994;72:429-445.
3. Murray CJL, Lopez AD. Assessing health needs: the Global Burden of Disease Study. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford Textbook of Public Health - Fourth edition. New York: Oxford University Press Inc.; 2004. p.243-54.
4. Pruss-Ustun A, Mathers C, Corvalan C, Woodward A. Introduction and methods: assessing the environmental burden of disease at national and local levels. WHO Environmental Burden of Disease Series, No. 1. Geneva: World Health Organization; 2003.
5. World Health Organization, Global Program on Evidence for Health Policy. National burden of disease studies: a practical guide. Geneva: World Health Organization; 2001.

Appendix

Table A1. Standard life expectancy table. Adapted from the World Health Organization (21)

	Sex			Sex			Sex	
Age	Males	Females	Age	Males	Females	Age	Males	Females
0	80,00	82,50	34	46,55	49,36	68	15,15	17,90
1	79,36	81,84	35	45,57	48,38	69	14,36	17,05
2	78,36	80,87	36	44,58	47,41	70	13,58	16,20
3	77,37	79,90	37	43,60	46,44	71	12,89	15,42
4	76,38	78,92	38	42,61	45,47	72	12,21	14,63
5	75,38	77,95	39	41,63	44,50	73	11,53	13,85
6	74,39	76,96	40	40,64	43,53	74	10,85	13,06
7	73,39	75,97	41	39,67	42,57	75	10,17	12,28
8	72,39	74,97	42	38,69	41,61	76	9,62	11,60
9	71,40	73,98	43	37,72	40,64	77	9,08	10,93
10	70,40	72,99	44	36,74	39,68	78	8,53	10,25
11	69,40	72,00	45	35,77	38,72	79	7,99	9,58
12	68,41	71,00	46	34,81	37,77	80	7,45	8,90
13	67,41	70,01	47	33,86	36,83	81	7,01	8,36
14	66,41	69,01	48	32,90	35,88	82	6,56	7,83
15	65,41	68,02	49	31,95	34,94	83	6,12	7,29
16	64,42	67,03	50	30,99	33,99	84	5,68	6,76
17	63,42	66,04	51	30,06	33,07	85	5,24	6,22
18	62,43	65,06	52	29,12	32,14	86	4,90	5,83
19	61,43	64,07	53	28,19	31,22	87	4,56	5,43
20	60,44	63,08	54	27,26	30,29	88	4,22	5,04
21	59,44	62,10	55	26,32	29,37	89	3,88	4,64
22	58,45	61,12	56	25,42	28,46	90	3,54	4,25
23	57,46	60,13	57	24,52	27,55	91	3,30	3,98
24	56,46	59,15	58	23,61	26,65	92	3,05	3,71
25	55,47	58,17	59	22,71	25,74	93	2,80	3,43
26	54,48	57,19	60	21,81	24,83	94	2,56	3,16
27	53,49	56,21	61	20,95	23,95	95	2,31	2,89
28	52,50	55,23	62	20,09	23,07	96	2,14	2,71
29	51,50	54,25	63	19,22	22,20	97	1,97	2,53
30	50,51	53,27	64	18,36	21,32	98	1,80	2,36
31	49,52	52,29	65	17,50	20,44	99	1,63	2,18
32	48,53	51,31	66	16,71	19,59	100	1,46	2,00
33	47,54	50,34	67	15,93	18,74			

Table A2. Average disability weights (ADW) for diseases and conditions including cancers and injuries. Adapted from the World Health Organization[21] (25,29)

Disease/Sequela	ADW	Disease/Sequela	ADW
Diseases and conditions except cancers and injuries			
TUBERCULOSIS		DIARRHOEAL DISEASES	
Cases	0.271	Diarrhoeal diseases – episodes	0.105
SYPHILIS		PERTUSSIS	
Congenital syphilis	0.315	Episodes	0.137
Primary	0.015	Encephalopathy	0.452
Secondary	0.048		
Tertiary - Neurologic	0.283	POLIOMYELITIS	
		Poliomyelitis - Cases - lameness	0.369
CHLAMYDIA			
Ophthalmia neonatorum	0.180	DIPHTHERIA	
Cervicitis	0.049	Episodes	0.231
Neonatal pneumonia	0.280	Neurological complications	0.078
Pelvic inflammatory disease	0.329	Myocarditis	0.323
Ectopic pregnancy	0.549		
Tubo-ovarian abscess	0.548	MEASLES - EPISODES	
Chronic pelvic pain	0.122	Measles - Episodes	0.152
Infertility	0.180		
Symptomatic urethritis	0.067	TETANUS	
Epididymitis	0.167	Episodes	0.638
GONORRHOEA		BACTERIAL MENINGITIS,	
Ophthalmia neonatorum	0.180	MENINGOCOCCAEMIA	
Corneal scar - Blindness	0.600	Streptococcus pneumoniae - Episodes	0.615
Corneal scar - Low vision	0.233	Haemophilus influenzae - Episodes	0.616
Cervicitis	0.049	Neisseria meningitidis - Episodes	0.615
		Meningococcaemia without	0.152
Pelvic inflammatory disease	0.169	meningitis – Episodes	
Ectopic pregnancy	0.549	Deafness	0.229
Tubo-ovarian abscess	0.548	Seizure disorder	0.100
Chronic pelvic pain	0.122	Motor deficit	0.381
Infertility	0.180	Mental retardation	0.459
Symptomatic urethritis	0.067		
Epididymitis	0.167	HEPATITIS B AND HEPATITIS C	
Stricture	0.151	Hepatitis B – Episodes	0.075
		Hepatitis C – Episodes	0.075
HIV			
Cases	0.135	MALARIA	
AIDS cases not on ART	0.505	Episodes	0.191
AIDS cases on ART	0.167	Anaemia	0.012
		Neurological sequelae	0.471

[21] Only average disability weights are given in this table. Many of sequela vary with age, and many vary also with treatment. For details, please see the special WHO documents available on the Internet (25), and a special WHO webpage on the Internet (20,26).

Table A2. Cont.

Disease/Sequela	ADW	Disease/Sequela	ADW
TRYPANOSOMIASIS		Cotemporaneous cognitive deficit	0.006
Episodes	0.350	Cognitive impairment	0.463
		Intestinal obstruction	0.024
CHAGAS DISEASE			
Infection	0.000	TRICHURIASIS	
Cardiomyopathy without congestive heart failure	0.062	High intensity infection	0.000
		Cotemporaneous cognitive deficit	0.006
Cardiomyopathy with congestive heart	0.270	Massive dysentery syndrome	0.116
Failure		Cognitive impairment	0.024
Megaviscera	0.240		
		HOOKWORM DISEASE -	
SCHISTOSOMIASIS		ANCYLOSTOMIASIS AND	
Infection	0.005	NECATORIASIS	
Advanced renal disease	0.104	High intensity infection	0.006
Advancerd hepatic disease	0.104	Anaemia	0.024
		Cognitive impairment	0.024
LEISHMANIASIS			
Visceral	0.243	LOWER RESPIRATORY	
Cutaneous	0.023	INFECTIONS	
		Episodes	0.279
LYMPHATIC FILARIASIS		Chronic sequelae	0.099
Hydrocele >15 cm	0.073		
Bancroftian lymphoedema	0.106	UPPER RESPIRATORY	
Brugian lymphoedema	0.116	INFECTIONS	
		Episodes	0.000
ONCHOCERCIASIS		Pharyngitis	0.070
Blindness	0.594		
Itching	0.068	OTITIS MEDIA	
Low vision	0.170	Episodes	0.023
		Deafness	0.229
LEPROSY			
Cases	0.000	MATERNAL HAEMORRHAGE	
Disabling leprosy	0.152	Episodes	0.000
		Severe anaemia	0.093
DENGUE			
Dengue fever	0.197	MATERNAL SEPSIS	
Dengue haemorrhagic fever	0.545	Episodes	0.000
		Infertility	0.180
JAPANESE ENCEPHALITIS			
Episodes	0.616	HYPERTENSIVE DISORDERS OF	
Cognitive impairment	0.468	PREGNANCY	
Neurological sequelae	0.379	Episodes	0.000
TRACHOMA		OBSTRUCTED LABOUR	
Blindness	0.581	Episodes	0.000
Low Vision	0.170	Stress incontinence	0.025
		Rectovaginal fistula	0.430
ASCARIASIS			
High intensity infection	0.000		

Table A2. Cont.

Disease/Sequela	ADW	Disease/Sequela	ADW
ABORTION		UNIPOLAR DEPRESSION	
Episodes	0.000	DISORDERD	
Infertility	0.180	Mild depressive episode	0.140
Reproductive tract infection	0.067	Moderate depressive episode	0.350
		Severe depressive episode	0.760
OTHER MATERNAL		Dysthymia	0.140
CONDITIONS			
Stress incontinence	0.025	BIPOLAR DISORDER	
		Cases	0.367
LOW BIRTH WEIGHT - ALL			
SEQUELAE		SCHIZOPHRENIA	
All sequelae	0.106	Cases	0.528
BIRTH ASPHYXIA AND BIRTH		EPILEPSY	
TRAUMA -		Cases	0.113
All sequelae	0.372		
		ALCOHOL USE DISORDERS	
PROTEIN-ENERGY		Cases	0.134
MALNUTRITION			
Wasting	0.053	ALZHEIMER AND OTHER	
Stunting	0.002	DEMENTIAS	
Developmental disability	0.024	Cases	0.666
IODINE DEFICIENCY		PARKINSON DISEASE	
Total goitre rate (G1 + G2)	0.000	Cases	0.351
Mild developmental disability	0.006		
Cretinoidism	0.255	MULTIPLE SCLEROSIS	
Cretinism	0.804	Cases	0.411
VITAMIN A DEFICIENCY		DRUG USE DISORDERS	
Xerophthalmia	0.000	Cases	0.252
Corneal scar	0.277		
		POST-TRAUMATIC STRESS	
IRON-DEFICIENCY ANAEMIA		DISORDER	
Mild	0.000	Cases	0.105
Moderate	0.011		
Severe	0.090	OBSESSIVE-COMPULSIVE	
Very severe	0.249	DISORDERS	
Cognitive impairment	0.024	Cases	0.127
DIABETES MELLITUS		PANIC DISORDER	
Cases	0.015	Cases	0.165
Diabetic foot	0.133		
Neuropathy	0.072	INSOMNIA (PRIMARY)	
Retinopathy - Blindness	0.552	Cases	0.100
Amputation	0.102		
		MIGRAINE	
		Cases	0.029

Table A2. Cont.

Disease/Sequela	ADW
MILD MENTAL RETARDATION ATTRIBUTABLE TO LEAD EXPOSURE	
Cases	0.361
GLAUCOMA	
Low vision	0.170
Blindness	0.600
CATARACTS	
Low vision	0.170
Blindness	0.570
HEARING LOSS, ADULT ONSET	
Mild	0.000
Moderate, untreated	0.120
Severe or profound, untreated	0.333
RHEUMATIC HEART DISEASE	
Cases	0.253
HYPERTENSIVE HEART DISEASE	
Cases	0.246
ISCHAEMIC HEART DISEASE	
Acute myocardial infarction	0.439
Angina pectoris	0.124
Congestive heart failure	0.201
CEREBROVASCULAR DISEASE	
First-ever stroke	0.920
Long-term stroke survivors	0.266
INFLAMMATORY HEART DISEASES	
All sequelae	0.252
COPD	
Mild and moderate symptomatic cases	0.170
Severe symptomatic cases	0.530
ASTHMA	
Cases	0.043
PEPTIC ULCER	
Cases not treated with antibiotic	0.042
CIRRHOSIS OF THE LIVER	
Symptomatic cases	0.330
APPENDICITIS	
Episodes	0.463
NEPHRITIS AND NEPHROSIS	
Acute glomerulonephritis	0.091
End-stage renal disease	0.098
BENIGN PROSTATIC HYPERTROPHY	
Symptomatic cases	0.038
SKIN DISEASES	
Cases	0.056
RHEUMATOID ARTHRITIS	
Cases	0.199
OSTEOARTHRITIS	
Hip	0.126
Knee	0.129
GOUT	
Cases	0.132
LOW BACK PAIN	
Episode of limiting low back pain	0.061
Episode of intervertebral disc displacement or herniation	0.061
Chronic intervertebral disc	0.121
CONGENITAL ANOMALIES	
Abdominal wall defect - Cases	0.850
Anencephaly - Cases	0.850
Anorectal atresia - Cases	0.850
Cleft lip - Cases	0.050
Cleft palate - Cases	0.103
Oesophageal atresia - Cases	0.850
Renal agenesis - Cases	0.850
Down syndrome - Cases	0.593
Congenital heart anomalies - Cases	0.323
Spina bifida - Cases	0.593
DENTAL DISEASES	
Dental caries - Cases	0.081
Periodontal disease - Cases	0.001
Edentulism - Cases	0.025

Table A2. Cont.

Disease/Sequela	ADW	Disease/Sequela	ADW
Malignant neoplasms			
CANCERS (SITES)		Corpus uteri	0.081
Mouth and oropharynx	0.118	Ovary	0.084
Oesophagus	0.217	Prostate	0.124
Stomach	0.217	Bladder	0.086
Colon and rectum	0.217	Lymphomas and multiple myeloma	0.073
Liver	0.239	Leukaemia	0.098
Pancreas	0.269		
Trachea, bronchus and lung	0.146	CANCERS - METASTASIS	0.750
Melanoma and other skin	0.045	STAGE (ALL SITES)	
Breast	0.078		
Cervix uteri	0.071	CANCERS TERMINAL (ALL SITES)	0.809
Injuries			
POISONING		INTERNAL INJURIES	
Short term	0.609	Short term	0.208
FRACTURES		OPEN WOUND	
Skull - Short term	0.431	Short term	0.108
Skull - Long term	0.384		
Face bones	0.223	INJURY TO EYES	
Vertebral column	0.266	Short term	0.108
		Long term	0.354
INJURED SPINAL CORD			
Injured spinal cord	0.725	AMPUTATIONS	
		Thumb	0.165
FRACTURES		Finger	0.102
Rib or sternum	0.199	Arm	0.308
Pelvis	0.247	Toe	0.102
Clavicle, scapula, or humerus	0.153	Foot	0.300
Radius or ulna	0.180	Leg	0.300
Hand bones	0.100		
Femur - Short term	0.372	CRUSHING	
Femur - Long term	0.272	Short term	0.218
Patella, tibia, or fibula	0.271		
Ankle	0.196	BURNS	
Foot bones	0.077	<20% - Short term	0.156
		<20% - Long term	0.002
DISLOCATIONS		>20% and <60% - Short term	0.469
Dislocated shoulder, elbow, or hip	0.074	>20% and <60% - Long term	0.255
Other dislocation	0.074	>60% - Short term	0.469
		>60% - Long term	0.255
SPRAINS			
Short term	0.064	INJURED NERVES	
		Short term	0.078
INTRACRANIAL INJURY		Long term	0.078
Short term	0.359		
Long term	0.359		

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Tests validity measures and analysis of the Receiver Operating Characteristic (ROC)**
Module: 2.31	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Jadranka Bozikov, PhD, Professor Andrija Stampar School of Public Health, School of Medicine, University of Zagreb, Croatia Lijana Zaletel Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Jadranka Bozikov Andrija Stampar School of Public Health Rockefellerova 4, 10000 Zagreb, Croatia E-mail: jbozikov@snz.hr
Keywords	Diagnostic tests, false negative rate, false positive rate, negative predictive value, positive predictive value, screening test, sensitivity, specificity, validity.
Learning objectives	After completing this module students should: • understand measures of tests validity and differences between them; • know how to calculate these measures and be capable to calculate them by themselves; • understand principles of ROC analysis, and how basic measures of tests validity are related to ROC analysis.
Abstract	Diagnosis is based on the results of diagnostic tests. Most of them are imperfect instruments, and make errors in both directions - a healthy individual can be classified as diseased, and vice versa. Ability of each diagnostic test to correctly classify patients as diseased or healthy is called validity of the test. The concept applies also in screening tests. There exist several nosological (sensitivity, specificity, etc.) and diagnostic measures (positive predictive value, negative predictive value) to assess validity of a test with a binary outcome. In other tests, ROC method could be used as a method of analysis.
Teaching methods	Teaching methods include introductory lectures, exercises, and interactive methods such as small group discussions. Students after introductory lectures first carefully read the recommended sources. Afterwards, they discuss the issue of tests validity measures. In continuation, they practice in groups of 2-3 students performing the procedure of calculation of all different measures of tests of validity using the programme tool (e.g. MS Excel) on given data. At the end, students compare and discuss their results.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple-choice questionnaire (MCQ) and case-study.

TESTS VALIDITY MEASURES AND ANALYSIS OF THE RECEIVER OPERATING CHARACTERISTIC (ROC)

Jadranka Bozikov, Lijana Zaletel-Kragelj

Theoretical background

Introduction

Diagnosis is based on the results of diagnostic tests (in the broadest meaning of that term). Most of these tests are imperfect instruments, and make errors in both directions - a healthy individual can be classified as diseased, and a diseased individual as healthy. Ability of each diagnostic test to correctly classify patients as diseased or healthy is called validity of the test. Assessing the validity of a test is especially important for the introduction of new diagnostic procedures (1-3).

The concept of test validity also applies in population studies, in the screenings of populations.

Test validity measures

Suppose that a diagnostic or screening test under observation provides us the binary outcome - the disease is present (usually referred to as "positive") or the disease is not present (usually referred to as "negative"). Therefore, "positive" means a greater probability of disease, whereas "negative" a greater probability of absence of disease (1-4). Of course, we are interested in:

- how well the patients with the disease are recognized by the test, or
- how well the test indicates whether the disease is really present.

For answering these two questions, it intuitively follows that, the test results are compared with the actual situation. Relation of results produced by some diagnostic to the actual state is presented in a 2×2 table of contingencies (Figure 1), also called the decision matrix.

Figure 1. Relation of results produced by a test to the actual state in the decision matrix

		Disease: Present	Disease: Absent
Test	Positive	a	b
	Negative	c	d

Box "a" represents the number of diseased examinees that are correctly recognized by the test as diseased. We say that the test recognized these examinees as **true positive** (TP). Box "b" represents number of healthy examinees that are incorrectly recognized by the test as diseased **- false positive** (FP). Box "c" represents number of diseased examinees that are incorrectly recognized by the test as healthy - **false negative** (FN). Box "d" represents number of healthy examinees that are correctly recognized by the test as healthy - **true negative** (TN). Figure 2 is presenting TP, FP, FN and TN classifications which are in fact absolute frequency measures of performance of a test.

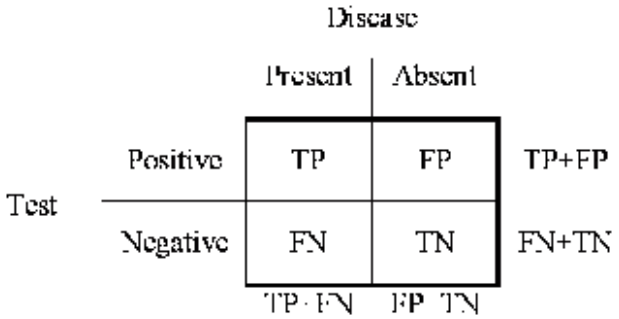
Figure 2. Relation of results produced by a test to the actual state with absolute measures of performance of a test in the decision matrix

		Disease: Present	Disease: Absent
Test	Positive	TP	FP
	Negative	FN	TN

Legend: TP – true positive classifications, FP – false positive classifications, FN – false negative classifications, TN – true negative classifications.

However, to answer the two questions raised above we need to form relative measures. Consequently, we need to complete the contingency table (the decision matrix) from Figure 1 and Figure 2 with marginal totals (Figure 3).

Figure 3. Complete decision matrix for calculating relative measures of performance of a test



To answer to the question "how well the patients with disease are recognized by the test", the so-called "nosological test validity measures" need to be formed, while to answer the question "how well the test indicates whether the disease is really present", the so-called "diagnostic test validity measures" need to be defined.

Nosological test validity measures and nosological probability

Nosological test validity measures are those validity measures where test results are compared to actual situation of the disease (1,2). There exist four nosological test validity measures:

- **sensitivity**, or true positive rate (TPR) – **nosological sensitivity** of test is proportion of sick people that test correctly recognizes as diseased (test-positive) of the total number of really diseased (Equation 1).

$$\text{sensitivity} = \frac{TP}{TP + FN} \qquad \textbf{Equation 1.}$$

- **specificity** or true negative rate (TNR) – **nosological specificity** of test is proportions of healthy people that test correctly identified of the total number of really healthy (Equation 2).

$$\text{specificity} = \frac{TN}{TN + FP} \qquad \textbf{Equation 2.}$$

Sensitivity and specificity are two main nosological measures of the validity of the test. The other two proportions are false positive and false negative ratio

- **false positive rate (FPR)** – false positive rate is a proportion of healthy that test incorrectly classified as diseased (Equation 3). False positive rate is at the same time 1-sensitivity (Equation 3).

$$FPR = \frac{FN}{TP + FN} = 1 - \text{sensitivity} \qquad \textbf{Equation 3.}$$

- **false negative rate (FNR)** – false negative rate is a proportion of the diseased that test wrongly placed as healthy (Equation 4). False negative rate is at the same time 1-specificity (Equation 4).

$$FNR = \frac{FN}{TP + FN} = 1 - \text{specificity} \qquad \textbf{Equation 4.}$$

All these equations could be seen also from the probability point of view. Let's mark events in following way:

- B means the presence of disease (the event "diseased").
- B' is the absence of disease (the event "not to be diseased" i.e., "healthy").
- indicates the presence of features (symptoms) or a positive test result.
- O' is absence of features (symptoms) or a negative test result.

From definitions for measures of validity of test and the concept of conditional probability (module 1.1.2) obviously sensitivity, specificity, FNR and FPR are (Equations 5-8):

$$\text{sensitivity} = P(O \mid B) = \frac{P(O \cap B)}{P(B)} \qquad \textbf{Equation 5.}$$

$$\text{specificity} = P(O'|B') = \frac{P(O' \cap B')}{P(B')}$$ **Equation 6.**

$$FNR = P(O'|B) = P(O|B)$$ **Equation 7.**

$$FPR = P(O|B') = P(O'|B')$$ **Equation 8.**

It should be noticed that it is necessary and sufficient to know two of four listed nosological probabilities (exactly two): for example, sensitivity, and specificity, because the other two are opposite probabilities.

Diagnostic test validity measures and diagnostic probability

Of course, medical doctors would be more interested in diagnostic test validity measures and hence diagnostic probabilities. Diagnostic test validity measures are those validity measures where actual situation of the disease is compared to test results (1,2). There exist two diagnostic test validity measures:

- **positive predictive value (PPV)** - positive predictive value represents proportion of really diseased of those who are positive on the test (Equation 9). PPV is known also as **diagnostic specificity**. This often causes confusion, especially because the attribute "diagnostic" often tends to be lost.

$$PPV = \frac{TP}{TP + FP}$$ **Equation 9.**

- **negative predictive value (NPV)** - negative predictive value represents proportion of real healthy individuals among individuals with negative test results. (Equation 10). NPV is known also as **diagnostic sensitivity**. Again, this often causes confusion, especially because the attribute "diagnostic" often tends to be lost.

$$NPV = \frac{TN}{FN + TN}$$ **Equation 10.**

From the probability point of view diagnostic probabilities are probabilities of presence/absence of disease for positive or negative test results. According to the concept of conditional probability and Bayes' theorem (module Probability - basic concepts) obviously PPV and NPV are (Equations 11 and 12):

$$PPV = P(B|O) = \frac{P(O|B) \cdot P(B)}{P(O|B) \times P(B) + P(O|B') \times P(B')}$$ **Equation 11.**

$$NPV = P(B'|O') = \frac{P(O'|B') \cdot P(B')}{P(O'|B) \times P(B) + P(O'|B') \times P(B')}$$ **Equation 12.**

These parameters are **influenced by disease prevalence in the observed population** (on nonconditional probabilities P(B)). Thus PPV (Equation 13) and NPV could be expressed also as (Equation 14):

$$P(B|O) = \frac{P(O|B) \times P(B)}{P(O|B) \times P(B) + P(O|B') \times P(B')} =$$ **Equation 13.**

$$= \frac{P(O|B) \times P(B)}{P(O|B) \times P(B) + (1 - P(O'|B')) \times (1 - P(B))} =$$

$$= \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1 - \text{specificity}) \times (1 - \text{prevalence})}$$

$$P(B'|O') = \frac{P(O'|B') \times P(B')}{P(O'|B') \times P(B') + P(O'|B) \times P(B)} =$$ **Equation14.**

$$= \frac{P(O'|B') \times P(B')}{P(O'|B') \times (1 - P(B)) + (1 - P(O|B)) \times P(B)} =$$

$$= \frac{\text{specificity} \times (1 - \text{prevalence})}{\text{specificity} \times (1 - \text{prevalence}) + (1 - \text{sensitivity}) \times \text{prevalence}}$$

Absolute and relative test accuracy

With test evaluation also go terms **absolute test accuracy (ATA)** (Equation 15) **and relative test accuracy (RTA)** (Equation 16), among which the first is and the second is not under the influence of disease prevalence in the tested population.

$$ATA = \frac{TP + TN}{TP + FP + FN + TN} \qquad \textbf{Equation 15.}$$

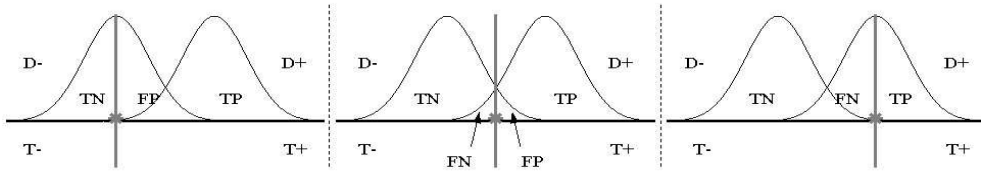
$$RTA = \frac{\left(\frac{TP}{TP + FP} + \frac{TN}{FN + TP}\right)}{2} \qquad \textbf{Equation 16.}$$

Mathematical basis of the above measures (which are used also for evaluation of expert systems) can be found in probability theory.

Receiver Operating Characteristic (ROC) analysis

However, many screening and diagnostic procedures in medicine does not only have two possible outputs - the disease is present or the disease is not present, but several values that can be measured on ordinal or even continuous scale of values. In such kind of diagnostic/screening test, we must first put the cut-off point on the scale of values in which to put the decision of a positive or negative result of test. In this cut-off point the decision matrix is constructed (3,5-7). Often it happens that we cannot just immediately put the best cut-off point. In this case we can put more cut-off points and in each of them we construct a decision matrix. In each, sensitivity (TPR), specificity (TNR), FNR and FPR are calculated. By varying the cut-off point these proportions change (Figure 4).

Figure 4. Change of the proportions of TP, TN, FP and FN test results as a function of changing the cut-off point for decision



LEGEND: * = the position of cut-off point, TN = true negative test results, TP = true positive test results, FN = false negative test results, FP = false positive test results, D- = disease not present, D+ = disease present, T- = negative test result, T+ = positive test result.

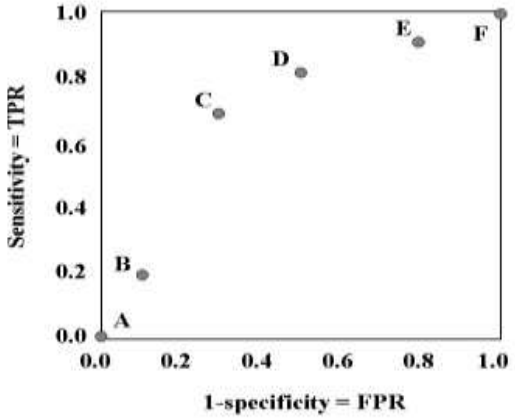
Choosing the best possible cut-off point depends on the cost-benefit associated with the classification of patients with the disease among those who do not have the disease, compared with the classification of healthy as those with the disease:

- when discovering the disease, which, if untreated, is mortal, we will gladly accept a slightly higher FPR, because we thus ensure that the TPR will be closer to 100%. But we must also consider the fact that people who would be labeled as diseased, but in fact they would be not, could suffer intangible costs (for example stress),
- in less serious diseases, or very expensive treatments, we would be interested in lowering the value of TPR on account of minimizing the FPR.

Analysis of sensitivity and specificity of the test, depending on setting borders is known as the **Receiver Operating Characteristic (ROC) analysis**. To deal with these multiple pairs of sensitivity and specificity values, we can draw a diagram using the **sensitivities as the y coordinates** and the **FPRs** (**1-specificities) as the x coordinates** (Figure 5) (3,5-7). The points in the ROC diagram could be fitted with a curve which can be smoothed (Figure 6).

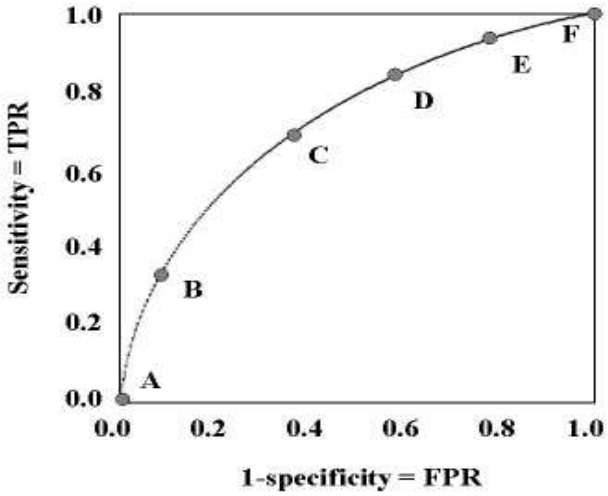
A good and very informative diagnostic/screening test is characterized by high sensitivity (TPR) values and low FPR values in all possible cut-off points. ROC curve of such a test is shifted to the left upper corner (Figure 7, curve W). A perfect test (Figure 7, curve W) has an area under the ROC curve of 1. By contrast, the ROC curve of poor diagnostic test is approaching the diagonal connecting the lower left corner of the image by right-upper one. On this diagonal (Figure 7, curve Z - diagonal), the sensitivity (TPR) values and FPR values in each cut-off points are the same.

Figure 5. Receiver Operating Characteristic (ROC) diagram made of four cut-off points (A-F).



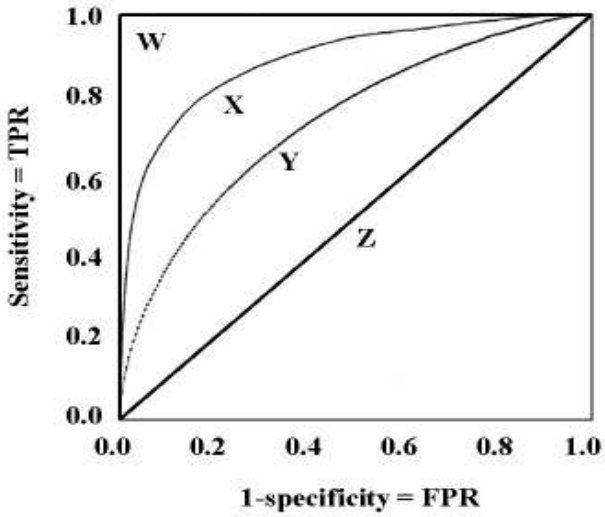
LEGEND: TPR = true positive rate, FPR = false positive rate.

Figure 6. Fitted and smoothed Receiver Operating Characteristic (ROC) curve



LEGEND: TPR = true positive rate, FPR = false positive rate.

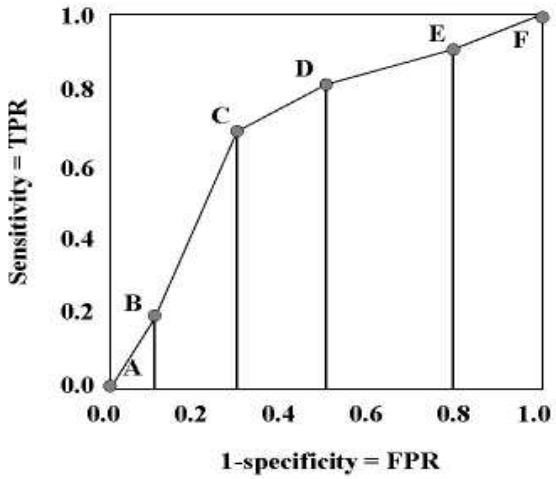
Figure 7. Fitted Receiver Operating Characteristic (ROC) curves for diagnostic tests, with varying degrees of informativity



LEGEND: (W = very informative test, X and Y = more/less moderately informative tests, Z = noninformative test).
TPR = true positive rate, FPR = false positive rate.

All this is reflected in the most important summary measure related to ROC curves – the **area under the curve** (AUC) (7,8). AUC is measuring the overall performance (accuracy, informativity) of a diagnostic test and is interpreted as the average value of sensitivity for all possible values of specificity. The AUC is calculated using different methods, the simplest being the trapezoid rule (7). In this process AUC is being divided into several parts, depending on how many measurement points we have (Figure 8).

Figure 8. Graphical representation as the basis of calculating area under Receiver Operating Characteristic (ROC) curve using trapezoid rule



All parts are added together at the end.

By using trapezoid rule, AUC, which is usually designated with greek letter θ, is in practice calculated as (Equation 17):

$$\theta = \theta_{1-0} + \theta_{2-1} + ... + \theta_{(i+1)-i} =$$
$$= \left[\left(FPR_1 - FPR_0 \right) \times \left(\frac{TPR_0 + TPR_1}{2} \right) \right] +$$
$$+ \left[\left(FPR_2 - FPR_1 \right) \times \left(\frac{TPR_1 + TPR_2}{2} \right) \right] + ...$$
$$... + \left[\left(FPR_{i+1} - FPR_i \right) \times \left(\frac{TPR_i + TPR_{i+1}}{2} \right) \right]$$

Equation 17.

For the case presented in Figure 8 the Equation 17 reads as (Equation 18):

$$\theta = \theta_{B-A} + \theta_{C-B} + ... + \theta_{F-E} =$$
$$= \left[\left(FPR_B - FPR_A \right) \times \left(\frac{TPR_A + TPR_B}{2} \right) \right] +$$
$$+ \left[\left(FPR_C - FPR_B \right) \times \left(\frac{TPR_B + TPR_C}{2} \right) \right] + ...$$
$$... + \left[\left(FPR_F - FPR_E \right) \times \left(\frac{TPR_E + TPR_F}{2} \right) \right]$$

Equation 18.

The problem of this method is that it gives an underestimates estimate of AUC. It should be noted that the more points there are, the better estimate of AUC we get. Much better estimate one can get by fitting the data to a binormal model with maximum likelihood estimates (7). However, this method is out of the scope of this module.

AUC can take any value between 0 and 1 (or 0 and 100% respectively). The closer AUC is to 1 (or 100%), the better the overall informativity (diagnostic performance) of the test. In another words, an area of 1 represents a perfect test. On the contrary an area of 0.5 represents a worthless test. This would be a test in which the positive test results were equally likely in both groups of people under investigation - those with the disease and those without it. A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system (9):

- AUC: 0.90-1.00 = excellent performance
- AUC: 0.80-0.90 = good performance
- AUC: 0.70-0.80 = fair performance
- AUC: 0.60-0.70 = poor performance
- AUC: 0.50-0.60 = fail performance

Case studies

Case study 1: test validity measures

Impact of disease prevalence on positive and negative predictive values

For illustration of the impact of disease prevalence in the population (sample) on PPV and NPV we will use two virtual sets of data representing two virtual situations. In both of them sensitivity and specificity of the test are exactly the same. Let's suppose that **sensitivity is 0.90** (90%) and **specificity is 0.95** (95%).

The first case is the case of open population in which the observed **prevalence of disease is 0,01** (1%). In Table 1 the situation in which sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%) is presented in details. The result of calculation of two nosological and two diagnostic measures of test validity for data set presented in Table 1 is presented in Table 2.

Table 1. The case of open population in which the observed prevalence of disease is 0.01 (1%), and sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%)

		Disease		
		Yes	**No**	**Total**
Test	**Positive**	90	495	585
	Negative	10	9405	9415
	Total	100	9900	10000

Table 2. The result of calculation of two nosological and two diagnostic measures of test validity in the case presented in Table 1

Measure	**Calculation**	**Result**
Prevalence	100/10000	0.010 (1.00%)
Sensitivity	90/100	0.900 (90.0%)
Specificity	9405/9900	0.950 (95.0%)
Positive predictive value	90/585	0.154 (15.4%)
Negative predictive value	9405/9415	0.999 (99.9%)

The second case is the case of hospital population where disease prevalence is 0.60 (60%). In Table 3 the situation in which sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%) is presented in detail.

Table 3. The case of hospital population in which the observed prevalence of disease is 0,60 (60%), and sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%)

		Disease		
		Yes	**No**	**Total**
Test	**Positive**	54	2	56
	Negative	6	38	44
	Total	60	40	100

Table 4. The result of calculation of two nosological and two diagnostic measures of test validity in the case presented in Table 3

Measure	**Calculation**	**Result**
Prevalence	60/100	0.600 (60.0%)
Sensitivity	54/60	0.900 (90.0%)
Specificity	38/40	0.950 (95.0%)
Positive predictive value	54/56	0.964 (96.4%)
Negative predictive value	38/44	0.864 (86.4%)

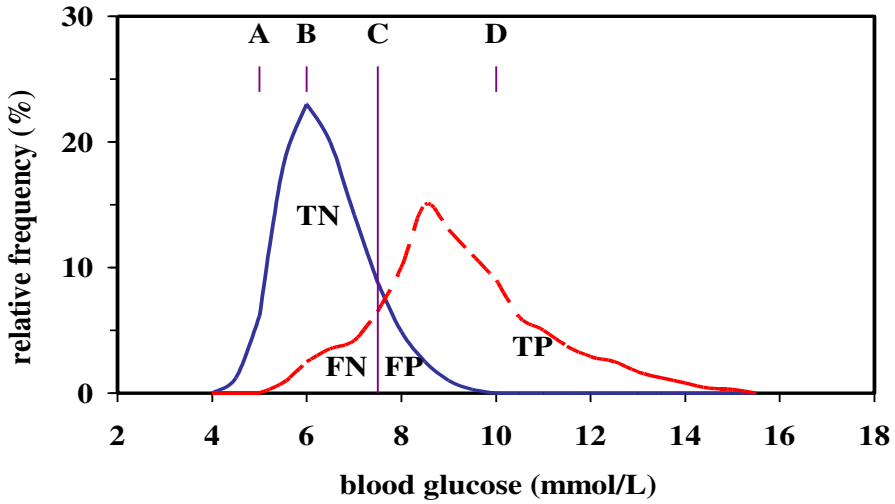
The result of calculation of two nosological and two diagnostic measures of test validity for data set presented in Table 3 is presented in Table 4. The comparison shows that the PPV in the second case is much, much higher than in the first case, although the sensitivity and specificity are exactly the same.

Case study 2: Receiver Operating Characteristic (ROC) analysis

The case of the test of blood glucose level

Curves in Figure 9. show the distribution of blood glucose level measured two hours after meal in healthy people and people suffering from diabetes.

Figure 9. The distribution of blood glucose level measured two hours after meal in healthy people (blue solid line) and people suffering from diabetes (red dashed line)

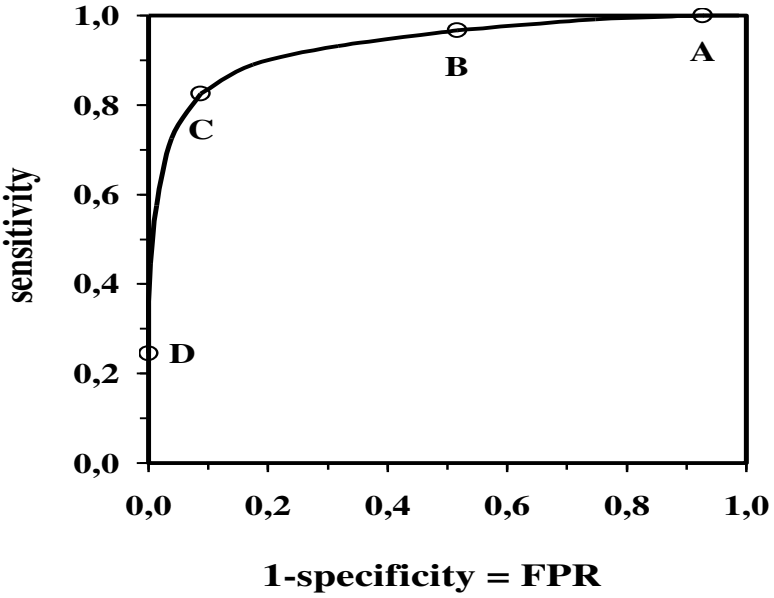


LEGEND: A, B, C and D = the position of four possible cut-off points, TN = true negative test results, TP = true positive test results, FN = false negative test results, FP = false positive test results.

Capital letters (A, B, C and D) in Figure 9 denote four possible cut-off points for final clinical decision. If, for example, in screening of population for diabetes limits are set very low (criterion A), the test will be very sensitive and detect all diseased people with very low specificity (many FP that request additional diagnostic tests). With increase of the limit (criteria) sensitivity will decrease, and specificity will increase (1-specificity or FPR will decrease).

In continuation a ROC diagram made of four cut-off points with fitted ROC curve (Figure 10) is constructed.

Figure 10. Fitted and smoothed Receiver Operating Characteristic (ROC) curve in the case of making decision in blood glucose level test

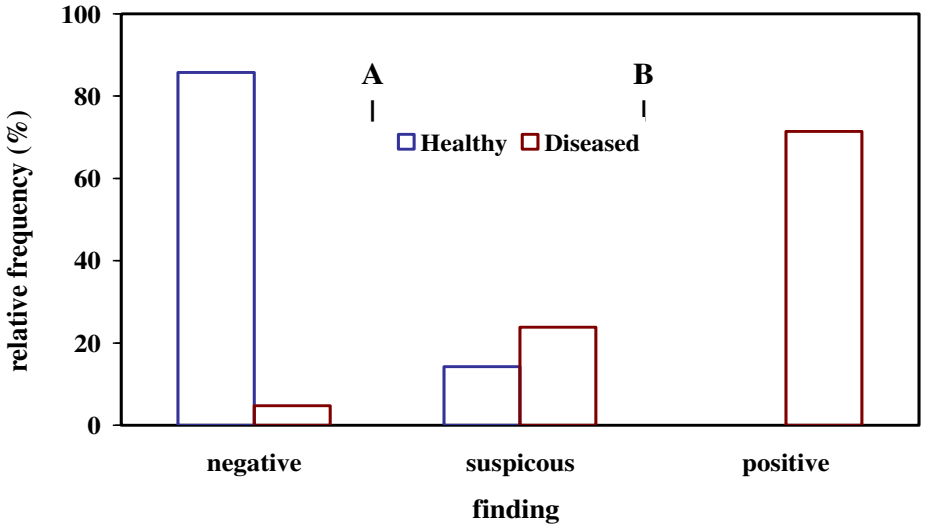


LEGEND: FPR = false positive rate.

The case of interpretation of radio cardiograms

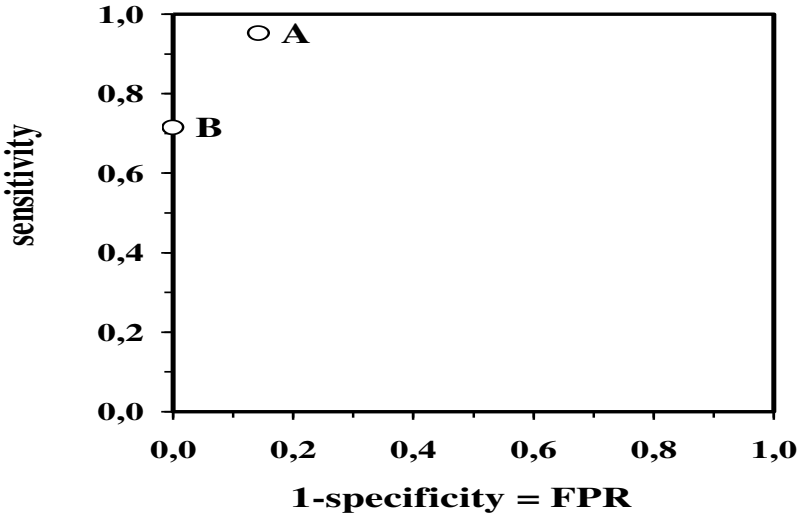
Of course, in qualitative diagnostic procedures, we can have only a few discrete values. Perfectly described example can be found in the work of Malcic dealing with classification of findings of radiocardiograms - the graphic record produced by radiocardiography (10).

Figure 11. The distribution of findings of radiocardiography that are classified as negative, suspect or positive in relation to the proven existence of shunt (10)



LEGEND: A, B = the position of two possible cut-off points.

Figure 12. Fitted and smoothed Receiver Operating Characteristic (ROC) curve in the case of classification of findings of radiocardiograms in diagnostic process of intracardial shunt from left to right (10)



LEGEND: FPR = false positive rate.

In the work of Malcic, the value of radiocardiographic testing in diagnostic process of intracardial shunt from left to right is test whose results are compared with results of catheterisation and angiography as the criteria for the existence of shunt in work (10). Figure 11 shows the distribution of findings of radiocardiography that are classified as negative, suspect or positive in relation to the proven existence of shunt. Capital letters (A and B) in Figure 11 denote two possible cut-off points for final clinical decision. Criteria A places all positive and suspect findings as positive and has a sensitivity somewhat lower than 100% (because two of 42 patients have negative radiocardiography findings) and the specificity is 71.4% (4/28 healthy have suspect radiocardiography finding). Criteria B only positive radiocardiography finding consider as positive and it is clearly evident that its sensitivity is 100% and specificity is 85.7% because some diseased examinees were not recognize as diseased.

In continuation, a ROC diagram made of four cut-off points with fitted ROC curve (Figure 12) is constructed.

Exercises

Task 1

In Table 5, a data set on performance of test T with binary results in diagnosing disease D is presented[22].

Table 5. Performance of test T with binary results in diagnosing disease D

		Disease D		
		Yes	**No**	**Total**
Test T	**Positive**	18	3	21
	Negative	5	214	219
	Total	23	217	230

Carefully read the theoretical part and calculate:

1. Nosological sensitivity of the test T
2. Nosological specificity of the test T
3. Diagnostic specificity of the test T
4. Diagnostic sensitivity of the test T

Define:

1. Another name for diagnostic specificity
2. Another name for diagnostic sensitivity

The results should be compared to the results of the teacher[23].

Task 2

In Table 6, a data set on performance of test T with several values of results in diagnosing disease D is presented.

Table 6. Performance of test T with several values of results in diagnosing disease D

Number of symptoms	Diseased	Healthy
1	12	48
2	13	30
3	40	20
4	60	10
5	40	2

Sketch the ROC curve (point) if the positive test is considered as the presence of three or more symptoms of the disease[24].

---

[22] This exercise is based upon real data presented in a paper on results in the test for antibody to neutrophil cytoplasmic antigens as a diagnostic aid for Wegener's granulomatosis (11).

[23] Answers Task 1:

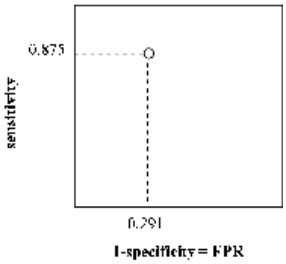
The results of calculating rates:

1. Nosological sensitivity:
   18/23 = 0.78 (78%)
2. Nosological specificity:
   214/217 = 0.99 (99%)
3. Diagnostic specificity:
   18/21 = 0.86 (86%)
4. Diagnostic sensitivity:
   214/219 = 0.98 (98%)

Replies to definitions:

1. Positive predictive value
2. Negative predictive value

[24] Answer Task 2:



Task 3

The group of researchers was interested how good is the performance of test T with several values of results in diagnosing disease D in three different groups of examinees.

Make the following:

1. From the data presented in Table 7 calculate three AUCs[25]. Hint: MS Excel could be very helpful in the calculation procedure.
2. Interpret calculated AUCs according to rough guide for classifying the accuracy of a diagnostic test.

The results should be compared to the results of the teacher[26].

Table 7. Performance of test T with several values of results in diagnosing disease D in three different groups of examinees

Cut-off point	Group 1		Group 2		Group 3	
	TPR	FPR	TPR	FPR	TPR	FPR
1	0.000	0.000	0.000	0.000	0.000	0.000
2	0.023	0.000	0.019	0.000	0.027	0.000
3	0.051	0.001	0.036	0.001	0.046	0.001
4	0.067	0.002	0.068	0.002	0.069	0.002
5	0.072	0.003	0.084	0.004	0.082	0.003
6	0.089	0.004	0.093	0.005	0.098	0.004
7	0.097	0.005	0.098	0.006	0.115	0.005
8	0.104	0.006	0.105	0.007	0.121	0.006
9	0.107	0.007	0.112	0.008	0.135	0.007
10	0.116	0.008	0.117	0.009	0.142	0.008
11	0.120	0.009	0.129	0.010	0.159	0.009
12	0.125	0.010	0.131	0.013	0.169	0.010
13	0.133	0.011	0.138	0.015	0.174	0.011
14	0.142	0.012	0.138	0.016	0.181	0.012
15	0.145	0.013	0.142	0.017	0.186	0.013
16	0.148	0.014	0.170	0.019	0.193	0.015
17	0.168	0.018	0.182	0.020	0.201	0.016
18	0.173	0.019	0.186	0.022	0.210	0.017
19	0.201	0.022	0.207	0.025	0.222	0.019
20	0.207	0.024	0.210	0.028	0.235	0.021
21	0.208	0.025	0.224	0.029	0.247	0.022
22	0.228	0.029	0.235	0.033	0.251	0.024
23	0.233	0.031	0.256	0.037	0.259	0.027
24	0.235	0.032	0.261	0.040	0.270	0.029
25	0.247	0.035	0.263	0.042	0.292	0.034
26	0.255	0.038	0.277	0.050	0.304	0.038
27	0.265	0.043	0.294	0.057	0.318	0.042
28	0.277	0.047	0.305	0.061	0.340	0.049
29	0.292	0.054	0.340	0.070	0.378	0.056
30	0.319	0.067	0.352	0.077	0.399	0.065
31	0.370	0.092	0.378	0.088	0.422	0.074
32	0.378	0.101	0.403	0.101	0.449	0.086
33	0.436	0.132	0.427	0.115	0.476	0.102
34	0.488	0.167	0.462	0.133	0.502	0.123
35	0.593	0.266	0.506	0.162	0.547	0.157
36	0.712	0.437	0.557	0.207	0.594	0.201
37	0.864	0.652	0.597	0.250	0.681	0.276
38	0.924	0.774	0.739	0.417	0.789	0.420
39	1.000	0.996	0.830	0.574	0.923	0.694
40	1.000	1.000	0.911	0.731	0.989	0.917
41			0.967	0.912	1.000	1.000
42			1.000	1.000		

[25] This exercise is based upon real data used in analysis of the effectiveness of multiple regression, discriminant analysis and logistic regression to identify and evaluate predictors of premature birth (12).

[26] Answers Task 3:

1. The results of calculating AUCs:
   Group 1: $\theta = 0.717$
   Group 2: $\theta = 0.732$
   Group 3: $\theta = 0.771$
2. Interpretation of the accuracy of a diagnostic test: in all three groups the performance of test is fair. However, the best is in Group 3.

References

1. Wulff HR. Klinicko prosudivanje (in Croatian). Zagreb: Jumena; 1982.
2. Bradley GW. Evaluation of tests. In: Bradley GW, editor. Disease, diagnosis and decisions. Chichester: John Wiley & Sons; 1993. p.70-90.
3. McNeil BJ, Keeler E, Adelstein JS. Primer on certain elements of medical decision making. N Engl J Med. 1975;293:211-215.
4. Ivankovic D. Osnove statisticke analize za medicinare. Zagreb: Medicinski fakultet Sveucilista u Zagrebu; 1989.
5. Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978;4:283-298.
6. Brismar J. Understanding receiver-operating-characteristic curves: a graphical approach. Am J Roentgenol 1991;157:1119-1121.
7. Kallin Westin L. Receiver operating characteristic (ROC) analysis: valuating discriminance effects among decision support systems. Umea: Deptartment of Computing Science, Umea University; 2001. Available from URL: http://www8.cs.umu.se/research/reports/2001/018/part1.pdf. Accessed: June 4, 2013.
8. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
9. Tape TG. Interpreting Diagnostic Tests. Omaha, NE: University of Nebraska Medical Center [Internet book]. Available from URL: http://gim.unmc.edu/dxtests/Default.htm. Accessed: June 4, 2013.
10. Malcic I. Dijagnosticka vrijednost radiokardiografije u djece s lijevo-desnim intrakardijalnim shuntom - ROC analiza (in Croatian). Lijec Vjesn. 1985;107:500-6.
11. Harrison DJ, Simpson R, Kharbanda R, Abernethy VE, Nimmo G. Antibodies to neutrophil cytoplasmic antigens in Wegener's granulomatosis and other conditions. Thorax. 1989;44:373-7.
12. Zaletel-Kragelj L. Analiza ucinkovitosti multiple regresije, diskriminantne analize in logisticne regresije pri identifikaciji in vrednotenju napovednih dejavnikov prezgodnjega poroda: doktorska disertacija (in Slovene). Ljubljana: Medicinska fakulteta; 1997.

Recommended readings

1. Bradley GW, editor. Disease, diagnosis and decisions. Chichester: John Wiley & Sons; 1993.
2. Kallin Westin L. Receiver operating characteristic (ROC) analysis: valuating discriminance effects among decision support systems. Umea: Deptartment of Computing Science, Umea University; 2001. Available from URL: http://www8.cs.umu.se/research/reports/2001/018/part1.pdf. Accessed: June 4, 2013.
3. Wulff HR. Klinicko prosudivanje (in Croatian). Zagreb: Jumena; 1982.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Screening**
Module: 2.32	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Mateja Krajc, MD, MSc Cancer Registry, Institute of Oncology, Ljubljana, Slovenia
Address for correspondence	Mateja Krajc, MD, MSc Cancer Registry, Institute of Oncology, Ljubljana Zaloska 2, Ljubljana, Slovenia E-mail: mkrajc@onko-i.si
Keywords	Breast cancer screening, cancer screening, screening, screening recommendations.
Learning objectives	After completing this module students should: • know what screenings are, what are their benefits and disadvantages, as well as key issues in screening; • be aware of ethical problems of screenings; • be able to list most important recommended screenings in Europe in different age groups; • be familiar with breast cancer screening process.
Abstract	There have been various definitions of screening over the years, but simply what we are talking about in screening is seeking to identify a disease or pre-disease condition in apparently healthy individuals. This concept is now widely accepted in most of the developed word. Used wisely, it can be a powerful tool in the prevention of a disease. Screening has important ethical differences from clinical practice as the health service is targeting apparently healthy people, offering to help individuals to make better informed choices about their health. The module is presenting basic theoretical background necessary for understanding the usefulness of screenings, the screening process, and potential risks, as well as it provides a case study of breast cancer screening.
Teaching methods	An introductory lecture gives the students a first insight into the characteristics of screenings. The theoretical knowledge is illustrated by a case study. After introductory lectures, students first carefully read the recommended readings. Afterwards, they discuss the characteristics of screenings, their benefits and disadvantages, as well as key issues in screening. They also discuss the basic criteria to be fulfilled before screening for any condition is introduced. In continuation, students are supposed to be more deeply engaged in the breast cancer screening process.
Specific recommendations for teachers	• ECTS: 0.2 • work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a computer room; • equipment: computers (one computer for 2-3 students), LCD projection equipment, internet connection, access to the bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire.

SCREENING*

Mateja Krajc

Theoretical background

Basic definitions and explanations of terms

Screening

According to the National Screening Committee of the United Kingdom Health Departments Second Report (1), screening is a public health service in which members of a defined population, who do not necessarily perceive they are at risk of, or are already affected by a disease or its complications, are asked a question or offered a test, to identify those individuals who are more likely to be helped than harmed by further tests or treatment to reduce the risk of a disease or its complications.

There have been various definitions of screening over the years (2-4) but, put simply, what we are talking about in screening is seeking to identify a disease or pre-disease condition in apparently healthy individuals. This concept is now widely accepted in most of the developed world. When used wisely, it can be a powerful tool in the prevention of a disease.

Screening has important ethical differences from clinical practice. The health service is targeting apparently healthy people, offering to help individuals to make better informed decisions about their health. Irrespective that screening has the potential to save lives or improve quality of life through early diagnosis of serious conditions it is not a fool-proof process. Screening can reduce the risk of developing a condition or its complications but it cannot offer a guarantee of protection. In any screening programme, there is an irreducible minimum of false positive results (wrongly reported as having the condition) and false negative results (wrongly reported as not having the condition).

Screening programmes

Screening programmes are public health services that are organized at the level of a large population and must be effectively monitored. Programmes must use research evidence to identify that they do more good than harm at a reasonable cost. Proposed new screening programmes should be assessed against a set of internationally recognised criteria. These criteria include the epidemiology of the condition, the screening test, any treatment options, and the acceptability of the screening programme.

The benefits of screening for disease prevention were first demonstrated in the 1940s, by the use of mass miniature radiography (MMR) for the identification of individuals with tuberculosis (TB). After the end of the Second World War, when effective treatment for TB was introduced, the use of MMR became widespread in many western countries. In 1968, WHO issued monograph Principles and Practice of Screening for Disease, which remains a landmark contribution to the screening literature.

Two basic types of screening

It is important to distinguish between **population based screening**, where people thought to be at risk are invited for screening, as in the national programmes for cancer of the breast and cervix, and **opportunistic screening** for prevention or case-finding, where individuals have sought medical advice for a specific symptom or complaint and opportunity is taken to suggest various other tests, such as the measurement of blood pressure or cholesterol, appropriate to their age and sex.

Criteria for screening

Before screening for any condition is introduced, the basic criteria have to be fulfilled (Table 1) (5). They are fundamental to the integrity of the screening process in any country.

The validity of screening test and the evaluation of screening

Validity

There are two measures to describe the validity of screening test – sensitivity and specificity. Both measures are conditional probabilities, and both are easy to understand using a decision matrix (Table 2) (6).

Sensitivity

Sensitivity (nosological) is defined as the ability of a test to detect all those with the disease in the screened population. This is expressed as the proportion of those with the disease in whom a screening test gives a positive result. Technically, it is a proportion of people with condition with positive test: a/(a+c) (Table 2).

* based on Holland WW, Stewart S. Screening in disease prevention: What works? Oxford: Radcliffe Publishing Ltd in association with the Nuffield Trust and the European Observatory on Health Systems amd Policies. 2005.

Table 1. Summary of criteria for screening (5)

Category	Criteria
Condition	The condition sought should be an important health problem whose natural history, including development from latent to declared disease, is adequately understood. The condition should have a detectable preclinical phase.
Target population	There should be a defined target population.
Diagnosis	There should be a suitable diagnostic test that is available, safe and acceptable to the population concerned. There should be an agreed policy, based on respectable test findings and national standards, as to whom to regard as patients, and the whole process should be a continuing one.
Treatment	There should be an accepted and established treatment or intervention for individuals identified as having the disease or pre-disease condition and facilities for treatment should be available.
Cost	The cost of case-finding (including diagnosis and treatment) should be economically balanced in relation to possible expenditure on medical care as a whole.
Screening test	Should be acceptable and safe.

Table 2. Decision matrix for derivation of the sensitivity and specificity of a screening test.

	Disease present	Disease absent	
Test positive	a	b	a+b
Test negative	c	d	c+d
	a+c	b+d	a+b+c+d

Specificity

Specificity is defined as the ability of a test to identify correctly those free of the disease in the screened population. This is expressed as a proportion of people free of the disease in whom the screening test gives a negative result. Technically, it is a proportion of people without condition with negative test: a/(a+c) (Table 2).

But one should be aware interpreting these measures since there are two kinds of sensitivity and specificity: nosological and diagnostic (7,8). So far we were speaking of nosological conditional probabilities. Other two important conditional probabilities are positive and negative predictive values (7,9).

Positive predictive value

Positive predictive value is the probability that a person with a negative test does not have the condition under screening. Technically, it is a proportion of people with positive test who have condition: a/(a+b) (Table 2). This measure is also known as diagnostic specificity.

Negative predictive value

Negative predictive value is the probability that a person with a negative test does not have the condition under screening. Technically, it is a proportion of people with negative test who do not have condition: d/(c+d) (Table 2). This measure is also known as diagnostic sensitivity.

All screening tests should aim to have high sensitivity and high specificity.

Evaluation

Evaluation must also be an integral part of any screening procedure. In 1971, Cochrane and Holland suggested seven criteria for evaluation and these remain as valid today as they were then (10) (Table 3).

Benefits and disadvantages

The benefits and disadvantages of screening have been fully described over the years and have been summarized by Chamberlain (11) (Table 4).

Table 3. Summary of criteria for evaluation of screening (10)

Factor	Criteria
Simplicity	The test should be simple to perform, easy to interpret and, where possible, capable of use by paramedics and other personnel.
Acceptability	Since participation in screening is voluntary, the test must be acceptable to those undergoing it.
Accuracy	The test must give a true measurement of the condition or symptom under investigation.
Cost	The expense of the test must be considered in relation to the benefits of early detection of the disease.
Repeatability	The test should give consistent results in repeated trials.
Sensitivity	The test should be capable of giving a positive finding when the individual being screened has the condition being sought.
Specificity	The test should be capable of giving a negative finding when the individual being screened does not have the condition being sought.

Table 4. Benefits and disadvantages of screening (11)

Benefits	Disadvantages
Improved prognosis for cases detected	Longer morbidity in cases where prognosis is unaltered
Less radical treatment which cures some early cases	Overtreatment of questionable abnormalities
Resource savings	Resource costs
Reassurance for those with negative test results	False reassurance for those with false-negative results Anxiety and sometimes morbidity for those with false positive results Hazard of screening test itself

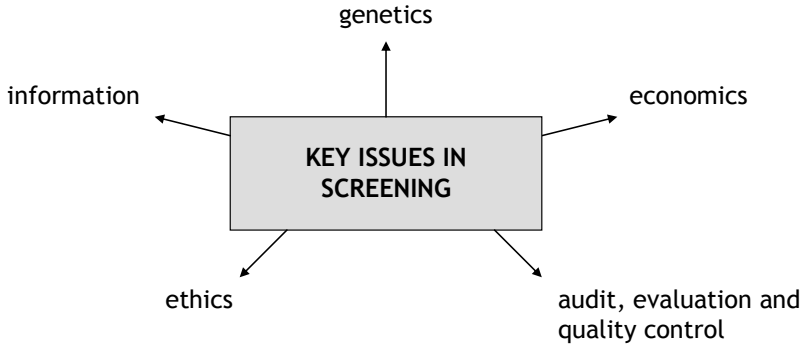
1. Benefits
   The benefits are very clear. Early and accurate diagnosis and intervention will lead to an improved prognosis in some patients. At this stage treatment may need to be less invasive.
2. Disadvantages
   The disadvantages are more complex. There will be longer periods of morbidity for patients whose prognosis is unchanged and there may be overtreatment of non-serious conditions or abnormalities identified. There are also resource costs in finding more illness both in terms of the tests themselves, the personnel costs and the subsequent management of whatever is found. There is the unpalatable certainty that some individuals with false-negative results will be given unfounded reassurance and that some with false positive results will experience, at the very least, unnecessary anxiety and, at the worst, inappropriate treatment.

Finally, there is the possibility, however remote, of hazard from the screening test itself. One point is particularly relevant here - there may be public demand (fuelled by vested interests) for the introduction of a screening test that does not meet the established criteria; an example of this is in screening for cancer of the prostate where the current screening test – prostate-specific antigen (PSA) – does not meet the criteria for accuracy or specificity.

Key issues in screening

There are a number of issues that are relevant at all stages and in every type of screening programme in any country, and are closely interrelated. There are five key issues in screening, being genetics, information, economics, ethics, and audit, evaluation and quality control (Figure 1).

Figure 1. Five key issues of screening



Before discussing the above mentioned key issues, one should consider components of an effectively organized screening programme. The components as described by Hakama (12) are as follows:

- The target population should be identified;
- Individuals in the population who are to be screened need to be identified;
- All those eligible for screening should be encouraged to attend – for example, by issuing a personal invitation, and offering suitable timing of screening examinations to suit the needs of those involved;
- There should be adequate premises, equipment and staff to ensure that the screening examination is done under pleasant circumstances and is acceptable to those attending;
- There should be an appropriate, satisfactory method of ensuring the maintenance of the best standards of the test(s) by:
  - initial and continuing training of the personnel conducting the test(s);
  - demonstration (by appropriate records) of the maintenance standards of equipment used in the examination – for example, calibration of X-ray machines in mammography;
  - routine checks of the validity of the tests performed – for example, random duplicate measurements for biochemistry, cytology, and reading of X-rays;
- There should be adequate and appropriate facilities for the diagnosis and treatment of any individual found to require this. There should be as little delay as possible between the screening attendance, advice that the screening test was negative, advice that the screening test result required further investigation, and referral to the appropriate centre for further investigation or treatment. A timetable should be established for these different procedures and there should be continuous monitoring to ensure that the time intervals between the various stages are complied with;
- There should be regular checks to ascertain the satisfaction level of those who have undergone the screening process – those investigated, the screen-negatives and those invited who have not participated;
- Finally, regular periodic checks should be made of the records of the screened individuals to ascertain their adequacy.

Genetics

In the last decade, genetic screening has developed very rapidly with the mapping of the human genome. Many see it as opening up a new era in the prevention, early diagnosis and identification of disease. However, caution is essential.

There are two objectives of screening for a recessive carrier state. One is to reduce the prevalence of the disorder and the other is to inform the reproductive choices of individuals and couples at risk. Information is thus regarded as worthwhile in itself, regardless of the possibility of prevention or treatment. While this type of screening can certainly help to evaluate risk and may be appropriate in certain high-risk groups. It should be carefully considered when to screen, if nothing can be done after the results of the screening test.

The main purpose of genetic screening at present is to prevent. In this it differs from much current screening practice and it must not be allowed to overlook the basic principles and criteria of screening.

Information

Information is another central concept in modern health care in general and also in screening. It must be provided in a correct way, so that possible participant may decide upon proper information, with the end-point being truly informed consent (or refusal) to participate.

Economics

As economic theory has entered the field, it has been recognized that screening may also do harm. All screening procedures involve the examination and testing of large numbers of individuals in order to find the few with an abnormality. There are two main consequences of this.

First, those who undergo screening are often understandably anxious while waiting for the result and become even more anxious if they have to undergo further investigation. Second, although most screening tests are simple, relatively cheap procedures in themselves, the actual costs are by no means trivial because of the large numbers involved.

Ethics

Any abnormality identified, whether in a national screening programme or in primary care, must be treatable and the investigation itself must not cause harm. Many believe that early diagnosis, particularly of cancer and heart disease, will lead to the possibility of treatment and improvement in prognosis. This is an attractive concept and can lead to a demand for a screening procedure to be introduced, irrespective of whether it has been shown that diagnosis guarantees an improved outcome.

Audit, evaluation and quality control

In any screening programme, as with any other service programme, adequate steps must be taken to ensure that the original objectives are being met and that the methodology meets appropriate standards.

The ideal method for evaluating a screening programme is the randomized controlled trial in which individuals in a population are allocated, at random, either to a group that is screened or to a group that receives only its normal medical care.

The components of an effectively organized screening programme have been described by Hakama (12), and have been already presented earlier in this module.

The importance of maintaining the quality of screening programmes should never be underestimated. Evaluation, audit and quality control should be an integral part of any screening programme to ensure that it is achieving what it has set out to do in a way that is acceptable to those involved.

The recommended screenings in Europe

There are several recommended screenings in Europe (3). They may be presented through different age groups, being:

- antenatal period;
- neonatal period;
- screening in childhood;
- screening in adolescence and early adulthood;
- screening in adults, and
- screening in eldery.

Recommended screenings by age groups

1. **Antenatal period**

   There are many routine screenings for the total population, and some screenings for high risk groups (Table 5). There are also some screenings under research review (Table 5) (12).

Table 5. Recommended screenings in antenatal period in Europe (13)

Condition under screening	Comment
Routine	
Anaemia	
Blood group and RhD status	
Hepatitis B	
HIV	
Risk factors for pre-eclampsia	
Rubella immunity	
Syphilis	
Asymptomatic bacteriuria	
Fetal anomalies: Anencephaly	
Spina bifida	
Chromosome abnormalities: Down syndrome	

High risk only
Thalassaemia/sickle cell disease
Tay-Sachs disease
Under research review
Duchenne muscular dystrophy
Chlamydia infection
Gestational diabetes
Fragile X syndrome
Hepatitis C
Genital herpes
HTLV1
Streptococcus B infection

2. Neonatal period

There are many routine screenings, and some screenings under research review (Table 6).

Table 6. Recommended screenings in neonatal period in Europe (13)

Condition under screening	Comment
Routine	
Phenylketonuria	Bloodspot
Congenital hypothyroidism	Bloodspot
Cystic fibrosis	Bloodspot
Sickle cell disease	Bloodspot
High risk only	
Congenital heart disease	Physical examination
Congenital cataract	Physical examination
Cryptorchism	Physical examination
Congenital dislocation of the hip/ developmental dysplasia of the hip test	Physical examination
Other congenital malformations	Physical examination
Hearing impairment	
Under research review	
Biotinidase deficiency	
Congenital adrenal hyperplasia	
Duchenne muscular dystrophya	

3. Childhood

Screenings, recommended in Europe in the childhood are presented in Table 7.

4. Adolescence and early adulthood

Screenings recommended in Europe in adolescence and early adulthood are presented in Table 8.

5. Adults

In Table 9, screenings, recommended in Europe in adulthood are presented.

Screening in adults is potentially big business. Media interest in health is insatiable, and anyone who reads the newspapers, watches television or listens to radio can hardly fail to be aware of the various diseases that may be lying in wait for them. Of course, it is of benefit if potential health problems can be identified early and treated. But society must beware of turning health into an obsession and must resist both the increasing medicalization of life and the growing politicization of medicine.

The national programmes for breast and cervical cancer should be continued but kept under review with an emphasis on quality control and on providing balanced and understandable information to enable women to make a truly informed choice without pressure from health professionals on whether or not to participate.

A national programme of screening for colorectal cancer by faecal occult blood testing in adults aged from 50 to 74 years has been agreed in the United Kingdom and on some other European countries but it is essential that adequate diagnostic, treatment and follow-up facilities are in place before it is introduced.

Screening for risk factors of coronary heart disease and stroke should be carried out in the primary care setting with advice, treatment and follow-up as appropriate. In the case of abdominal aortic aneurysm, it now seems clear that ultrasound screening in men aged 65 years and over would reduce mortality from this condition, although the benefit for those aged over 75 years has been questioned. As with colorectal cancer, however, national implementation should await the certainty that adequate facilities and resources are available. In the case of screening for diabetic retinopathy, close attention must be paid to audit and the need to be absolutely clear about how, when and where to screen.

Table 7. Recommended screenings in childhood in Europe (13)

Condition under screening	Comment
Hearing impairment	• Follow-up on neonatal programme where indicated • School entry "sweep" test to continue • Case-finding to identify late onset or progressive impairment • Investigation of any children with educational or behavioural problems
Amblyopia and impaired vision	• Orthoptist screening in 4–5-year-olds • Attention to be paid to children who miss this test for any reason
Dental disease	• School dental screening mandatory and should continue, but should be kept under research review • Early contact with dentists to be encouraged • Problems include shortage of dentists and lack of parental compliance, especially among the more deprived
Congenital hip dysplasia/ developmental dysplasia of the hip (CHD/DDH)	• Children identified by neonatal screening to be reviewed • Parental observations and concerns to be investigated
Deprived, disadvantaged or socially isolated children	• Need to identify such children and instigate screening/case-finding where relevant

Table 8. Recommended screenings in adolescence and early adulthood in Europe (13)

Condition under screening	Comment
Chlamydia	• Opportunistic screening of those aged 25 and under who access sexual health services or primary care.

6. Elderly

Society is facing a major challenge in how best to maintain health and quality of life in populations where the proportion of people aged over 60 years now outnumbers those aged under 16 and the number of individuals aged over 85 is rising.

A system of regular surveillance and case-finding in primary care would seem to be the most appropriate form of screening, particularly in those aged 75 and over, but the resource implications of this must be confronted. Several simple tests, such as identifying difficulties with sight or hearing or problems with feet, can make a huge difference to the comfort and quality of life. Depression is another area where identification and treatment could improve well-being. Social and community support are also vital in enabling older people to enjoy as independent and contented a life as possible. The emphasis in screening at this stage of life should be on improving quality of life and preserving function and independence, rather than on providing "heroic" treatments to prevent mortality.

In Table 10, screenings, recommended in Europe in elderly people are presented.

Table 9. Recommended screenings in adulthood in Europe (3)

Condition under screening	Comment
Breast cancer	• National programme should be continued but kept under close review with emphasis on quality control, staff training and good information
Cervical cancer	• National programme should be continued with review of alternative types of tests and of age range of those eligible and frequency of screening, • Good information to be a priority
Colorectal cancer	• National screening programme by faecal occult blood testing for adults aged 50–74 years
Abdominal aortic	• Ultrasound screening of men aged 65 and over Aneurysm seems a reasonable proposition provided the necessary resources are in place
Diabetic retinopathy	• National programme of screening for all diabetics aged over 12. It is essential to be quite clear about how, when and where screening should happen to ensure effective implementation
Risk factors for coronary heart disease	• Weight surveillance/case-finding approach in primary care
(CHD)/stroke	
Blood pressure	
Cholesterol	
Smoking cessation	

Table 10. Recommended screenings in elderly in Europe (3)

Condition under screening	Comment
Hypertension	Physical assessment
Early heart failure	Physical assessment
Hearing loss	Physical assessment
Vision loss	Physical assessment
Incontinence	Physical assessment
Lack of physical activity	Physical assessment
Foot problems	Physical assessment
Review of medication	Physical assessment
Depression	Mental assessment
Alcohol use	Mental assessment
Falls	Social assessment
Undernutrition	Social assessment
Isolation	Social assessment

Conclusions

Screening programmes and practices vary widely across the countries of the European Union (EU). This is inevitable given the differing structures and financing of health services, and differing demographic features of the population. There are, however, key objectives to strive for.

These include having one national body per country responsible for practice and policy, scrupulous adherence to the long-established screening criteria, accurate population registers, greater uniformity of

access across different mareas of a given country and across different socioeconomic groups, and sound research evidence on which to base practice. The wide variation in practice in Europe illustrates the complexity of screening. Some lessons, however, stand out. Key points of screening in the EU are (3):

- Antenatal screening programmes for Down syndrome and spina bifida are performed only in a few countries and are mainly optional. They are often only recommended to women at high risk.
- Neonatal screening for phenylketonuria is systematically recommended in all countries belonging to the EU before May 2004, except Finland.
- Breast cancer screening and cervical cancer screening programmes are recommended in some European countries.
- HIV screening is more common among the new Member States and three Candidate Countries and covers specific vulnerable groups, such as pregnant women and blood donors.
- TB screening is performed in a few European countries, especially central and eastern European countries, such as Hungary, Romania and Turkey.
- Not all the countries follow the basic criteria for screening. A population register to allow recall and follow-up of patients is often missing. A single national body for reviewing tests and practice is rare.

Case study

Breast cancer screening

Cancer screenings

At present the following screening tests meet requirements for organized screening programmes (Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC) OJ L 327/34-38) (14):

- pap smear screening for cervical abnormalities starting at the latest by the age of 30 and definitely not before the age of 20,
- mammography screening for breast cancer in women aged 50-69 in accordance with European guidelines on quality assurance in mammography,
- faecal occult blood screening for colorectal cancer in men and women age 50-74.

Decisions on implementation of cancer screening programmes must be made as part of a general priority-setting exercise on the use of healthcare resources (15-17).

Other cancer screening test are not yet recommended for EU-wide population-based cancer screening, although they already may be used in individual screening on demand. Such tests may provide individual benefits but at the same time may also lead to adverse effects for individuals (e.g. unfounded anxiety) and the public (e.g. additional financial burden). Recommendations for such tests cannot be made until they have shown to have benefits such as reducing disease-specific mortality or improving survival (18-20).

Potentially promising screening tests currently being evaluated in randomised controlled trials, include:

- prostate-specific antigen (PSA) testing for prostate cancer;
- mammography screening for women aged 40-49 for breast cancer;
- immunological Faecal Occult Blood Testing (FOBT) for colorectal cancer;
- flexible colonoscopy for colorectal cancer.

Once the effectiveness of a new screening test has been demonstrated, evaluation of modified testing methods may be possible using intermediate/surrogate endpoints, if the positive predictive value of such endpoints is sufficiently established. Some examples of screening methods which fall into this category are listed below:

- any novel alternative tests for faecal occult blood;
- liquid-based cervical cytology;
- testing for high risk human papilloma virus (HPV) infection;
- other novel methods for the preparation or interpretation of cervical specimens.

Any screening test which has been demonstrated to be effective should be offered on a population basis only in organised screening programmes, with quality assurance at all levels and full information about the benefits and risks (21,22).

Breast cancer screening

Breast cancer is currently the most frequent cancer and the most frequent cause of cancer induced deaths in women in Europe. Demographic trends indicate a continuing increase in this substantial public health problem. Systematic early detection through screening, effective diagnostic pathways and optimal treatment have the ability to substantially lower current breast cancer mortality rates and reduce the burden of this disease in the population.

In order that these benefits may be obtained, high quality services are essential. These may be achieved through the underlying basic principles of training, specialisation, volume levels, multidisciplinary team working, the use of set targets and performance indicators and audit. Ethically these principles should be regarded as applying equally to symptomatic diagnostic services and screening.

The primary aim of a breast screening programme is to reduce mortality from breast cancer through early detection. Unnecessary workup of lesions which show clearly benign features should be avoided in order to minimise anxiety and maintain a streamlined cost-effective service. Women attending a symptomatic breast service have different needs and anxieties and therefore mixing of screening and symptomatic women in clinics should be avoided.

Foundamental points and principles of the European guidelines for quality assurance of breast cancer screening programmes

Fundamental points and principles of the 4th edition of the European guidelines for quality assurance of breast cancer screening programmes are (23):

- breast cancer screening is a complex multidisciplinary undertaking, the objective of which is to reduce mortality and morbidity from the disease without adversely affecting the health status of participants. It requires trained and experienced professionals using up-to-date and specialised equipment;
- screening usually involves a healthy and asymptomatic population which requires adequate information presented in an appropriate and unbiased manner in order to allow a fully informed choice as to whether to attend. Information provided must be balanced, honest, adequate, truthful, evidence-based, accessible, respectful and tailored to individual needs where possible (23-25);
- mammography remains the cornerstone of population-based breast cancer screening. Due attention must be paid to the requisite quality required for its performance and interpretation, in order to optimise benefits, lower mortality and provide an adequate balance of sensitivity and specificity;
- physico-technical quality control must ascertain that the equipment used performs at a constant high quality level providing sufficient diagnostic information to be able to detect breast cancer using as low a radiation dose as is reasonably achievable. Routine performance of basic test procedures and dose measurements is essential for assuring high quality mammography and comparison between centres;
- full-field digital mammography can achieve high image quality and is likely to become established due to multiple advantages such as image manipulation and transmission, data display and future technological developments. Extensive clinical, comparative and logistical evaluations are underway;
- the role of the radiographer is central to producing high quality mammograms which, in turn, are crucial for the early diagnosis of breast cancer. Correct positioning of the breast on the standard lateral oblique and cranio-caudal views is necessary to allow maximum visualisation of the breast tissue, reduce recalls for technical inadequacies and maximise the cancer detection rate;
- radiologists take prime responsibility for mammographic image quality and diagnostic interpretation. They must understand the risks and benefits of breast cancer screening and the dangers of inadequately trained staff and sub-optimal equipment. For quality loop purposes the radiologist performing the screen reading should also be involved at assessment of screen detected abnormalities;
- all units carrying out screening, diagnosis or assessment must work to agreed protocols forming part of a local quality assurance (QA) manual, based on national or European documents containing accepted clinical standards and published values. They should work within a specialist framework, adhering to set performance indicators and targets. Variations of practices and healthcare environments throughout the member states must not interfere with the achievement of these;
- a robust and reliable system of accreditation is required for screening and symptomatic units, so that women, purchasers and planners of healthcare services can identify those breast clinics and units which are operating to a satisfactory standard. Any accreditation system should only recognise centres that employ sufficiently skilled and trained personnel;
- the provision of rapid diagnostic clinics where skilled multidisciplinary advice and investigation can be provided is advantageous for women with significant breast problems in order to avoid unnecessary delay in outline of management planning or to permit immediate discharge of women with normal/benign disease;
- population breast screening programmes should ideally be based within or closely associated with a specialised breast unit and share the services of trained expert personnel.

Key performance indicators for monitoring in population based breast cancer screening

Key performance indicators to be monitored in any population based breast cancer screening programme are presented in Table 11.

Table 11. Summary table of key performance indicators to be monitored in any population based breast cancer screening programme

	Performance indicator	Acceptable level	Desirable level
1.	Target optical density	1.4 - 1.9 OD	1.4 - 1.9 OD
2.	Spatial resolution	> 12 lp/mm	> 15 lp/mm
3.	Glandular dose – PMMA thickness at 4.5 cm	< 2.5 mGy	< 2.0 mGy
4.	Threshold contrast visibility	< 1.5%	< 1.5%
5.	Proportion of women invited that attend for screening	> 70%	> 75%
6.	Proportion of eligible women reinvited within the specified screening interval	> 95%	100%
7.	Proportion of eligible women reinvited within the specified screening interval + 6 months	> 98%	100%
8.	Proportion of women with a radiographically acceptable screening examination	97%	> 97%
9.	Proportion of women informed of procedure and time scale of receiving results	100%	100%
10.	Proportion of women undergoing a technical repeat screening examination	< 3%	< 1%
11.	Proportion of women undergoing additional imaging at the time of the screening examination in order to further clarify the mammographic appearances	< 5%	< 1%
12.	Proportion of women recalled for further assessment		
	• initial screening examinations	< 7%	< 5%
	• subsequent screening examinations	< 5%	< 3%
13.	Proportion of screened women subjected to early recall following diagnostic assessment	< 1%	0%
14.	Breast cancer detection rate, expressed as a multiple of the underlying, expected, breast cancer incidence rate in the absence of screening (IR):		
	• initial screening examinations	3 x IR	> 3 x IR
	• subsequent screening examinations	1.5 x IR	> 1.5 x IR
15.	Interval cancer rate as a proportion of the underlying, expected, breast cancer incidence rate in the absence of screening:		
	• within the first year (0-11 months)	30%	< 30%
	• within the second year (12-23 months)	50%	< 50%
16.	Proportion of screen-detected cancers that are invasive	90%	80-90%
17.	Proportion of screen-detected cancers that are stage II+:		
	• initial screening examinations	NA	< 30%
	• subsequent-regular screening examinations	25%	< 25%
18.	Proportion of invasive screen-detected cancers that are node-negative:		
	• initial screening examinations	NA	> 70%
	• subsequent-regular screening examinations	75%	> 75%
19.	Proportion of invasive screen-detected cancers that are ≤ 10 mm in size		
	• initial screening examinations	NA	≥ 25%
	• subsequent-regular screening examinations	≥ 25%	≥ 30%
20.	Proportion of invasive screen-detected cancers that are < 15 mm in size	50%	> 50%
21.	Proportion of invasive screen-detected cancers < 10 mm in size for which there was no frozen section	95%	> 95%
22.	Absolute sensitivity of FNAC	> 60%	> 70%
23.	Complete sensitivity of FNAC	> 80%	> 90%

24.	Specificity of FNAC	> 55%	> 65%
25.	Absolute sensitivity of core biopsy	> 70%	> 80%
26.	Complete sensitivity of core biopsy	> 80%	> 90%
27.	Specificity of core biopsy	> 75%	> 85%
28.	Proportion of localised impalpable lesions successfully excised at the first operation	> 90%	> 95%
29.	Proportion of image-guided FNAC procedures with insufficient result	< 25%	< 15%
30.	Proportion of image-guided FNAC procedures from lesions subsequently proven to be malignant, with an insufficient result	< 10%	< 5%
31.	Proportion of patients subsequently proven to have breast cancer with a pre-operative FNAC or core biopsy at the diagnosis of cancer	90%	> 90%
32.	Proportion of patients subsequently proven to have clinically occult breast cancer with a pre-operative FNAC or core biopsy that is diagnostic for cancer	70%	> 70%
33.	Proportion of image-guided core/vacuum procedures 0o0with an insufficient result	< 20%	< 10%
34.	Benign to malignant open surgical biopsy ratio in women at initial and subsequent examinations	≤ 1:2	≤ 1:4
35.	Proportion of wires placed within 1 cm of an impalpable lesion prior to excision	90%	> 90%
36.	Proportion of benign diagnostic biopsies on impalpable lesions weighing less than 30 grams	90%	> 90%
37.	Proportion of patients where a repeat operation is needed after incomplete excision	10%	< 10%
38.	Time (in working days) between:		
	• screening mammography and result	15 wd	10 wd
	• symptomatic mammography and result	5 wd	
	• result of screening mammography and offered assessment	5 wd	3 wd
	• result of diagnostic mammography and offered assessment	5 wd	
	• assessment and issuing of results	5 wd	
	• decision to operate and date offered for surgery	15 wd	10 wd
39.	Time (in working days) between:		
	• screening mammography and result		
	≤ 15 wd	95%	> 95%
	≤ 10 wd	90%	> 90%
	• symptomatic mammography and result		
	≤ 5 wd	90%	> 90%
	• result of screening mammography and offered assessment		
	≤ 5 wd	90%	> 90%
	≤ 3 wd	70%	> 70%
	• result of symptomatic mammography and offered assessment		
	≤ 5 wd	90%	> 90%
	• assessment and issuing of results		
	≤ 5 wd	90%	> 90%
	• decision to operate and date offered for surgery		
	≤ 15 wd	90%	> 90%
	≤ 10 wd	70%	> 70%

LEGEND: OD=optical density, PMMA=test object material (polymethylmethacrylate), IR=incidence rate, NA=not applicable, FNAC=fine needle aspiration citology, wd= week days

Exercises

Task 1

Carefully read the theoretical background of this module, and recommended readings.

Task 2

Critically discuss the differences between population based and oportunistic screening.

Task 3

Name the basic criteria to be fulfilled before screening for any condition is introduced.

Task 4

How do we describe the validity of screening test? Provide an example.

Task 5

List some advantages and disadvantages of the screening.

Task 6

Which screening tests for cancer meet all requirements for organized screening programmes?

Task 7

Critically assess the advantages and disadvantages of a population based breast cancer screening programme.

References

1. Health Departments of the United Kingdom. Second Report of the National Screening Committee, October. London: Health Departments of the United Kingdom, 2000.
2. Last JM. A dictionary of epidemiology. Oxford, Oxford University Press, 2001.
3. Holland WW, Stewart S, Masseria C. Policy Brief. Screening in Europe. Brussels: WHO Regional Office for Europe and European Observatory on Health Systems and Policies, 2006. http://www.euro.who.int/document/e88698.pdf.
4. Wilson JMG, Jungner G. Principles and practice of screening for disease. Geneva: World Health Organization, 1968. http://whqlibdoc.who.int/php/WHO_PHP_34.pdf.
5. Wilson JMG, Jungner G. Principles and practice of screening for disease. Geneva: WHO; 1968. Available from: http://www.who.int/bulletin/volumes/86/4/07-050112BP.pdf.
6. Detels R, McEwen J, Beaglehole R, Tanaka H, eds. Oxford textbook of public health. Oxford: Oxford University Press, 2002.
7. Wulff HR. Diagnosis and principles of probability [in Croatian]. In: Wulff HR. Klinicko prosuđivanje. Zagreb: Jugoslavenska medicinska naklada, 1979. p. 73-91.
8. Adamic S. Probability and clinical decision making [in Slovene]. In: Temelji biostatistike. Ljubljana: Medicinska fakulteta Univerze v Ljubljani, 1989. p.123-132.
9. Bradley GW. Evaluation of tests. In: Bradley GW (editor). Disease, diagnosis and decisions. Chichester: John Wiley & Sons, 1993. p. 70-90.
10. Cochrane AL, Holland WW.Validation of screening procedures. British Medical Bulletin. 1971;27(1):3-8.
11. Chamberlain Jocelyn M. Which prospective screening programmes are worthwhile? Journal of epidemiology and Community health. 1984;38:270-277.
12. Hakama M. Screening for cancer. Scandinavian Journal of Social Medicine Supplement. 1086; 37:17-25.
13. http://www.healthforallchildren.co.uk, http://nelh.nhs.uk/screening/child.
14. Council Recommendation of 2 December 2003 on cancer screening (2003/878/EC) OJ L 327/34-38.
15. Grimes DA, Schulz KF. Uses and abuses of screening tests. The Lancet. 2002;359:881-884. http://www.hawaii.edu/cm/assets/images/Lancet-research.pdf.
16. Primic Zakelj M, Zadnik V, Pogacnik A, Ursic Vrscaj M. Presejanje za raka maternicnega vratu v Sloveniji in drzavni program ZORA. Radiol Oncol. 2006;40(Suppl 1):S143-S148.
17. Stabuc B (urednik). XIII. Seminar "In memoriam dr. Dusana Reje". Zgodnje odkrivanje raka – izzivi za prihodnost. Ljubljana: Zveza slovenskih druđtev za boj proti raku, 2005.
18. Premik M. Prvine vrednotenja presejalnih programov. Slov pediatr. 1996;3:36-40.
19. Premik M. Vrednotenje presejalnih programov. Zdrav. vars. 1998;37:442-446.

20. Oblak C. Zgodnje odkrivanje, zdravljenje in spremljanje bolnikov z rakom prostate. Radiol Oncol. 2006;40(Suppl 1):S115-S120.
21. Novak Mlakar D, Maucec Zakotnik J. Presejanje za raka debelega crevesa in danke pri asimptomatski populaciji v starosti 50-69 let. Radiol Oncol. 2006;40(Suppl 1):S95-S106.
22. U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. U.S. Preventive Services Task Force (USPSTF). http://www.ahrq.gov/clinic/uspstfix.htm#Recommendations.
23. Perry N, Broeders M, de Wolf C, Tornberg S, Holland R, von Karsa L, Puthaar E (eds). European Guidelines for Quality Assurance in Breast Cancer Screening and Diagnosis. Fourth Edition. European Commission. Luxembourg, Office for Official Publications of the European Communities, 2006 (ISBN 92-79-01258-4).
24. Maucec Zakotnik J ,Krajc M, Primic Zakelj M. Vloga in delovanje preventivnih/presejalnih centrov za dojke. Radiol Oncol. 2006;40(Suppl 1):S31-S44.
25. Primic Zakelj M, Krajc M, Maucec Zakotnik J. Presejanje za raka dojk – kje smo danes. Radiol Oncol. 2006;40(Suppl 1):S25-S30.

Recommended readings

1. Grimes DA, Schulz KF. Uses and abuses of screening tests. The Lancet. 2002;359:881-884. http://www.hawaii.edu/cm/assets/images/Lancet-research.pdf.
2. Holland WW, Stewart S, Masseria C. Policy Brief. Screening in Europe. Brussels: WHO Regional Office for Europe and European Observatory on Health Systems and Policies, 2006. http://www.euro.who.int/document/e88698.pdf.
3. CINDI Slovenija. http://www.cindi-slo.net.
4. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC, 1999.
5. U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. U.S. Preventive Services Task Force (USPSTF). http://www.ahrq.gov/clinic/uspstfix.htm#Recommendations.
6. Wilson JMG, Jungner G. Principles and practice of screening for disease. Geneva: World Health Organization, 1968. http://whqlibdoc.who.int/php/WHO_PHP_34.pdf.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Rapid assessment and response**
Module: 2.33	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Enida Imamovic, DDM, MPH Public Health Institute of Federation BH, Bosnia and Herzegovina Dragana Niksic, MD, PhD, Professor Medical Faculty, University of Sarajevo, Bosnia and Herzegovina
Address for Correspondence	Enida Imamovic Public Health Institute of Federation BH Titova 9, 71000 Sarajevo, Bosnia and Herzegovina Tel: + 387 33 56 46 36, Fax: + 387 33 56 46 02 E-mail: enidai(@bih.net.ba; e.imamovic@zzizfbih.ba
Keywords	Rapid assessment, response.
Learning objectives	After completing this module students and public health professionals should: • understand the specificities of the RAR methodology; • be able to identify the priority problem in their community/country which could be assessed by RAR; • be able to identify the key community representatives whose support is crucial for RAR implementation.
Abstract	Rapid assessment, response (RAR) is a methodological approach used in situations where data is needed extremely quickly, where official statistics or other research techniques cannot provide up-to-date data about a particular public health issue. RAR combines different methods (observations, mapping, pools, focus groups and interviews) and sources of data. RAR can be focused on a particular disease (e.g. HIV/AIDS), a particular population group (e.g. the Roma people), or on a particular risk behaviour (e.g. smoking).
Teaching methods	After the introductory lecture and the recommended literature readings, students (all together) have to identify a key public health problem in their community/country that could be assessed by RAR. After that, students work in small groups; each group has the task to develop one of the RAR methods related to the chosen public health problem. Then, the groups will hold presentations where will discuss possible interventions, too. The result of this team work is drafting up a RAR Protocol for a chosen priority.
Specific recommendations for teachers	• work under teacher supervision: 30%, individual students' work proportion: 70%; • facilities: a lecture room, a computer room, rooms for small-group work; • equipment: one computer for five students, access to the Internet and bibliographic data-bases; • training materials: recommended readings and RAR frame Protocol • target audience: master degree students according to Bologna scheme.
Assessment of Students	Multiple-choice questionnaire (MCQ); group presentations: RAR protocol

RAPID ASSESSMENT AND RESPONSE

Enida Imamovic, Dragana Niksic

Theoretical background

Introduction

Since Rapid Assessment and Response (RAR) represents a relatively new methodological approach, a whole range of guides has been prepared in order to describe how to assess and respond rapidly to a broad variety of public health issues.

Numerous international experts worked hard on development of different RAR guides, such as the technical guide (TG-RAR), guides for injecting drug use (IDU-RAR), sexual behaviour guides (SEX-RAR), guides aimed to assess risk behaviour among young persons that are particularly prone to risk (EVYP-RAR) and other (1-7).

The RAR is used in situations where data is needed extremely quickly, in situations where official statistics or other research techniques cannot provide up-to-date data. The main objective of the RAR is achieved if the effective intervention measures are implemented (8-13).

The approach is a part of a significant preventive strategy which provides information, skills and interventions aimed to decrease risk behaviour within different population groups.

Radical socio-economic changes including still-existing war consequences in the South-Eastern Europe (SEE) countries have contributed to an increase of risk behaviour especially among the youth for which there was no official data. Therefore, UNICEF enabled the data collection by applying the RAR. As part of the project "HIV/AIDS prevention among the young people of South-Eastern European countries" RAR was carried out parallel in five SEE countries: Albania, Bosnia and Herzegovina, Macedonia, Croatia and the Federal Republic of Yugoslavia (14). The main aim was to gain a rapid insight into the risk behaviour of young people and adopt evidence based intervention measures.

Methodology

It is possible to carry out a research applying Rapid Assessment and Response both in wider and narrower geographic areas, i.e. on the local community level, in certain towns and countries, or even on a larger scale – on the regional level.

Advantages of the RAR methodology are reflected in an active participation of the local community, use of existing information and different methods which are equally important and in the possibility to analyze data on several levels (Box 1) (15-21).

RAR can be focused on:

- a disease or state (HIV/AIDS, sexually transmitted diseases – STD, Tuberculosis, pregnancy, etc.),
- a specific population group (e.g. a vulnerable group of young people aged 10-24, refugees, the Roma people, prisoners, victims of violence, etc.),
- places where the problem occurs (institutions designated to provide accommodation and upbringing of minors, prisons or a part of the local community at risk),
- risk behaviour (sexual habits, consummation of alcohol, tobacco smoking, drugs, etc).

The research provides new data that cannot be attained in another way (primary data). In order to collect the primary data a variety of methods is used: observation, mapping, poll, focus group and interview (1-5,13,15,16).

All the existing data (secondary data) attained from official institutions, governmental and non-governmental organizations, international organizations and data that was collected from another relevant research is used as well.

Advisory body

Key representatives of the community are involved in the research by forming an advisory body. This contributes to the success of the RAR implementation. Representatives of all sectors and target groups (government and the non-government sectors, media, health professionals, teachers, parents, religious communities) can contribute to the rapid assessment and make easier the direct implementation of intervention measures thereby contributing to the sustainability of ideas and practices advocated by RAR (22).

Research team

Representatives of different administrative levels and sectors make the research team. The research team consists of public health experts, representatives of the non-governmental sector as well as the members of the target groups. That is important because the target groups are often unwilling to enclose

personal information to the public sector. The first task of the core research team is to define the general and specific research goals, key assessment areas, working methods, the time frame and budget.

Rapid assessment methods

RAR enables the collection of quantitative and qualitative data:

- quantitative data gives us an insight into the distribution of the phenomenon and grant statistical confirmation. This data is vital for the planning and implementation of the intervention,
- qualitative data gives us deeper insight into the complexity of the examined phenomenon (e.g. risk behaviour) and into their influence on the examined group (young people aged 15-24 years). This data is necessary for the specification of goals and intervention design (1-5,18,20).

Collection of existing or secondary data

The secondary data does not have to be directly related to research goals but is related to the research framework. The sources of secondary data are different and those are: demographic, vital and health statistical data. All of this data is relevant, whether it is published or not (project reports of governmental and non-governmental organizations, scientific periodicals, books, and some other publications) as well as the data derived from different media (media files). It is important not to waste precious time on collecting secondary data that does not have a significant value in the end.

Collection of primary data

Primary data is collected from all available sources, such as:

- "ordinary" population,
- members of population groups falling within the scope of the research,
- members of target risk groups, including former members,
- key-informants (street vendors, tenants, parents, social workers, teachers, policemen, drug dealers, procurers, etc.).

Observations

Simple observing, listening and recording used as a research method, gains us an insight into first-hand information which usually remains "hidden" but is directly related to behaviour, relations and contexts of risk.

Mapping

The mapping method is applied simultaneously with the observation method. It is used to illustrate the aspects of the environment in which the persons live and where they practice certain behaviour. The application of the mapping method assists us to identify the areas where the risk behaviour is "more prominent", which also provides the opportunity for targeted interventions to take place. However, one should bear in mind that the research is completely anonymous and that by applying the mapping method we do not wish to compromise the confidentiality of the data collected.

Polls

Polls are used as another method of data collection. The research instrument consists of obligatory questions which, for the sake of standardization, have to be identical for all the countries in which RAR is implemented. Obligatory questions are posed to all the respondents to which the research refers (e.g. young people aged 14-24), that means to all the participants in the poll, the interview and the focus group. Apart from the obligatory questions, in order to provide a better insight into the specificity of on their own territory, each country could define a set of additional questions. Prior to field work, there is an obligatory pre-testing, after which a definite version of the questionnaire is created.

Focus groups

Focus groups provide us with a better insight into the group dynamics and the social norms related to the behaviour of the respondents. Focus groups give us information that cannot be obtained by interviews or other methods. Specially prepared questions are used for different focus groups.

The focus group method is applicable for getting information from different groups' respondents and that can contribute to a better understanding of problems and answer the questions which cannot be answered in another way (e.g. students, parents, teachers, pedagogues, health services, drug takers, nongovernmental organizations representatives, the governmental sector, etc).

Interview

Interviews attempt to get replies to questions which are too personal to answer in a poll or a focus group. Interviews are used to get information from individuals belonging to risk population groups, health workers, teachers, politicians and other relevant representatives of the community.

Data collection techniques

Since RAR collects both qualitative and quantitative data it uses the "snow ball" strategy. Once we start getting the same data by applying a certain method, or working with a certain group of respondents, the collection of data by applying that method is to be interrupted. Therefore, data is collected until the saturation moment has been reached. The "snow ball" strategy pays attention to questions that had previously not been answered clearly enough.

Ethical issues

Throughout the research, great emphasis is laid on ethical issues. Regardless of the method applied for data collection respondents are asked for permission for their replies to be recorded. The participation in the RAR being completely voluntary, each respondent can at any time refuse to participate or refuse to reply to any of the questions posed.

With the aim to collect data on the spot, all the working team members have accreditation cards and a "letter of presentation" which serves to identify them and to explain the general goals of the RAR. The research being completely anonymous, the confidentiality of data is by no means an issue, and even the detailed maps and observations are used for research exclusively.

Assessment areas

Data is being collected and analyzed in terms of the assessment areas:

- context (the entire environment),
- risk and protective behaviour,
- health and social consequences,
- existing interventions.

The context assessment enables data on factors that are related to health risk situations as well as on factors important for intervention measures implementation.

The assessment of risk and protective behaviour enables data on the risk type and its extent on one side, and existing protection measures (prevention) on the other. Preventive measures decrease the risk that a certain phenomena can bear among the population.

Data on negative health and social consequences of risk behaviour is collected within the assessment of social and health consequences, and within the framework of intervention assessment. In order to assess the necessity of future interventions, data on all existing intervention measures is collected.

Data analysis

Data entry and the statistical analysis of all data is an ongoing process. Questions that were not answered remain "empty" throughout the data entry process. Questionnaires filled in by persons whose age does not correspond to recommended age groups are not taken into account. Data is analyzed in a descriptive way, which implies the use of frequencies, averages and percentages.

The results attained by qualitative methods are first entered into team members' journals during field work and are later transferred into specially designed activity tables. All information obtained by different methods is entered into a separate activity table. Key responses from all activity tables are entered into summary tables or mega-tables for all four assessment areas: context, risk and protective behaviour, health and social consequences and existing interventions.

Triangulation

Same question data received from different methods is triangulated. This ensures a high data validity. Definition of triangulation: comparing or crosschecking data that was collected using different methods (e.g. focus group, existing information, polls) and from different data sources (e.g. injecting drug users, health service providers, politicians).

Data that has not been confirmed by triangulation cannot be taken into account in the final analysis.

Induction

A majority of research questions is defined prior to RAR, while others can be posed during the data collection process. Induction is an approach according to which a hypothesis is stated and then data is collected – data that will confirm, deny or modify that hypothesis. This can be repeated several times with a rapid assessment.

Results dissemination

The RAR results are presented in a form of Report. There are different ways of results' dissemination (through media, meetings with decision-makers, health service providers, non-governmental organizations, donors, etc.) and it is best to use as many different methods as possible. A quick dissemination of results to all community representatives (politicians, media, donors, health and non-health professionals, the NG sector, groups of respondents, etc.) is of great significance. This phase is particularly important for the intervention measures adoption and implementation.

Intervention

First of all, the community's response to a phenomenon which is being examined has to be rapid. Several alternatives have to be thought of as intervention measures. These alternatives have to be acceptable to decision-makers, donors and the institutions in charge for the implementation. The level of intervention can be individual, group or it can be structural (in terms of legislative change). Individual intervention is less efficient because an individual is perceived as a "victim". Increasing the knowledge of individuals about the problem does not have a significant effect because other requirements for the over-all rehabilitation (social, professional, economic requirements) have not been met. Group and structural interventions have a more comprehensive effect as they are creating requirements for sustainable community changes (18,20-22).

Conclusions

Rapid Assessment and Response (RAR) is a relatively new methodological approach. Its main characteristics are a rapid preparation, rapid adoption and implementation of intervention measures (Box 1). This method can be carried out on a local level (municipality, town, country), or a sub-regional and regional level. RAR is applicable in situations where data is needed extremely quickly, where official statistics or other research techniques cannot provide relevant data.

The engagement of relevant representatives of the community in the research from the beginning contributes to the overall success of the research. As a result of a successfully conducted RAR effective intervention measures are implemented to a wide spectrum of public and health issues (HIV/AIDS and STDs; consummation of alcohol, tobacco and drugs; sexual and reproductive health, violence, injuries and accidents; nutrition and mental health, etc.).

Box 1. Rapid assessment and response (RAR) features

Rapid Assessment and Response – RAR

- a rapid preparation and implementation,
- collect new data only when the existing data are insufficient and unreliable,
- crosschecking data derived from different sources (triangulation),
- complexity of approach; social, cultural, religious, political and economic context,
- engagement and support by the decision-maker,
- engagement of the community in all phases of research,
- a rapid and all-encompassing intervention.

Exercises

Task 1

Improve knowledge about the RAR methodology and its application by reading the recommended literature and browsing the Internet. Criticize and discuss the application of this method with colleagues.

Task 2

Find out if this method had already been applied in the country in which the students live or in the neighbouring countries and in which area had the research been carried out, then read the final reports.

Task 3

Select a significant municipality/national/regional public health problem for the RAR research (taking into account that it has to be a problem for which relevant, updated data is not available) which calls for a rapid assessment and intervention.

Task 4

Final paper (group work): Create a RAR research protocol for a chosen priority problem (Box 2).

Box 2. Guide for preparation of the rapid assessment and response (RAR) protocol

RAR PROTOCOL (example)	
Title	RAR on risk behaviour among youth
Problem formulation	A situational analysis pointed out at the significance of the public health problem (risk behaviour among youth) and the necessity for a rapid assessment of the situation.
Research topic	Define the target population (e.g. young people aged 15-24), risk behaviour related to the disease (e.g. HIV/AIDS) or state (e.g. pregnancy), territory (e.g. municipality, country)
The overall goal	Rapid assessment of risk behaviour among the youth and adoption of intervention measures with the aim to improve their health

Specific goals	A decrease in risk behaviour among youth (smoking, consummation of alcohol, *drugs, increase condom use…)*
Methodology	Secondary data (specify) Primary data (specify the methods) Timeframe (set) Key participants (specify)
Data base, data input and data processing	Define the data base (which), data input (when and who), data processing (when and who)
Data analysis	Simulate the analysis in terms of: • Context: socio-economic and political environment • Risk and protective behaviour • Health and social consequences • Existing interventions
Results	Specify the anticipated results (e.g. Final RAR Report adopted, intervention measures implemented...)
Dissemination	Simulate result dissemination: to whom (the key participants, decision makers, population...) and how (methods-*media, meetings, round tables…)*
Resources	Simulate the necessary human and material resources
RAR team	Advisory body (specify the members according to their occupation, profession and institution they work for) Research team members (their number in terms of their profession and role: team leader and deputy, members of the working group, poll-takers – field work, etc.
Timeframe	Set in terms of working phases (organigramme)
Budget	Simulate the framework budget

References

1. World Health Organisation. Rapid Assessment and Response Technical Guide (TG-RAR). WHO/HIV/2002. Available from URL: http://www.who.int/docstore/hiv/Core_All/Core_All.html. Accessed: June 03, 2013.
2. WHO, Regional Office for Europe. The RAR guide on injecting drug use: Social Change and Mental Health. 1998.
3. UNAIDS and WHO. The RAR guide on substance use and sexual risky behaviour. 1998.
4. World Health Organization. The Rapid Assessment and Response Guide on Psychoactive Substance Use and Sexual Risk Behaviour (SEX-RAR Guide). Geneva, 2002. ISBN 9241545585. Available from URL: http://www.who.int/mental_health/media/en/686.pdf. Accessed: June 04, 2013.
5. UNICEF. Rapid Assessment and Response on HIV/AIDS among Especially Vulnerable Young People in South Eastern Europe. 2002. Available from URL: ttp://www.cpha.ca/uploads/progs/infectious/rarunicf.pdf. Accessed: June 04, 2013
6. World Health Organisation. Rapid assessment and response Adaptation guide for work with especially vulnerable young people (RAR-EVYP). 2004. Available from URL: http://who.int/hiv/pub/prev_care/en/youngpeoplerar.pdf. Accessed: June 04, 2013.
7. World Health Organisation, Regional Office for Europe. The RAR guide on injecting drug use (IDU-RAR), Social Change and Mental Health. Draft for field testing. Copenhagen, 1998. Available from URL: http://gametlibrary.worldbank.org/FILES/1329_IDU%20Rapid%20Assessment%20and%20Response%20Guidelines.pdf. Accessed: June 04, 2013.
8. UNICEF. A decade of transition. MONEE Project, SEE/CIS/Baltic. 2000.
9. UNAIDS. Global strategy Framework on young people and HIV/AID. 2000.
10. UNICEF/Area Balkans. HIV/AIDS in South Eastern Europe. 2001.
11. Needle RH, Trotter RT II et all. Rapid Assessment of the HIV/AIDS Crisis in Racial and Ethnic Minority Communities: An Approach for Timely Community Interventions. Am J Pub Health 2003;93:6970-6979.

12. Rhodes T, Stimson GV, Fitch C, Ball A, Renton A. Rapid assessment, injecting drug use and public health. Lancet 1999;354:65-68.
13. Kamineni VV, Turk T, Wilson N, Satyanarayana S, Chauhan LS. A rapid assessment and response approach to review and enhance advocacy, communication and social mobilisation for tuberculosis control in Odisha state, India. BMC Public Health. 2011 Jun 10;11:463.
14. UNICEF/CPHA. Rapid Assessment and Response (RAR) for HIV/AIDS in South Eastern Europe. Available from URL: http://www.cpha.ca/en/activities/balkans.aspx#2. Accessed: June 04, 2013.
15. Trotter RT II, Needle RR, Goosby E, Bates C, Singer M. A methodological model for rapid assessment, response, and evaluation: the RARE program in public health. Field Methods.2001;13:137-159.
16. McNall M, Foster-Fishman PG. Methods of rapid evaluation, assessment, and appraisal. Am J Evaluat. 2007;28:151-68.
17. Brown D R, Hernandez A, Saint-Jean G, et alI. A participatory action research pilot study of urban health disparities using rapid assessment response and evaluation. Am J Public Health. 2008;98:28 38.
18. Pope C. Qualitative Research: Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ. 1995;311:42-45.
19. Siri P, Inchaurraga S. 'First steps': using rapid assessment and response methods to develop research, intervention and advocacy capacity for addressing drug use in Rosario City, Argentina. Int J Drug Policy. 2000 Mar 1;11(1-2):125-132.
20. Needle RR, Trotter RT II, Goosby E, Bates C, Von Zinkernagel D. Methodologically sound rapid assessment and response: providing timely data for policy development on drug use interventions. Int J Drug Policy. 2000;11:19-23.
21. Stimson GV, Fitch C., Jarlais DD, Poznyak V, Perlis T, Oppenheimer E, Rhodes T. The Who Phase II Drug injection collaborative rapid assessment and response studies of Injection Drug Use: Knowledge gain, capacity building, and intervention development in a multisite study. Am J Public Health. 2006;96:288-95.
22. Stanton B, Kim N, Galbraith J, Parrot M. Design Issues Addressed in Published Evaluations of Adolescents HN-Risk Reduction Interventions. J Adolesc Health 1996. 18:387-396.

Recommended readings

1. World Health Organisation. Rapid Assessment and Response Technical Guide (TG-RAR). WHO/HIV/2002. Available from URL: http://www.who.int/docstore/hiv/Core_All/Core_All.html. Accessed: June 03, 2013.
2. World Health Organization, Regional Office for Europe. The RAR guide on injecting drug use: Social change and mental health. copenhagen: World Health Organization, Regional Office for Europe; 1998. Available from URL: http://www.unodc.org/documents/hiv-aids/IDU%20rapid%20ass.%20and%20resp.%20guide.pdf. Accessed: June 04, 2013.
3. World Health Organization. The Rapid Assessment and Response Guide on Psychoactive Substance Use and Sexual Risk Behaviour (SEX-RAR Guide). Geneva, 2002. Available from URL: http://www.who.int/mental_health/media/en/686.pdf. Accessed: June 04, 2013.
4. UNICEF. Rapid Assessment and Response on HIV/AIDS among Especially Vulnerable Young People in South Eastern Europe. 2002. Available from URL: http://www.cpha.ca/uploads/progs/infectious/rarunicf.pdf. Accessed: June 04, 2013.

HEALTH INVESTIGATION: ANALYSIS – PLANNING – EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Introduction to health technology assessment**
Module: 2.34	ETCS (suggested): 0.25
Author(s), institution(s)	Jelena Marinkovic, Prof. BM, PhD The author is professor at the School of Medicine at Belgrade University.
Address for correspondence	Institute of Medical Statistics and Informatics, School of Medicine & School of Public Health and Health Management, University of Belgrade, Dr Subotica 15, 11000 Belgrade, Serbia; Email: marinkovic.j@gmail.com
Keywords	Decentralization/centralization, decision-making in health purchasing, evidence-based knowledge, health technology assessment, outcomes research, policy and practice, quality of care.
Learning objectives	The basic premise of this module is that the purpose of assessment of a particular technology (including elements of technology itself, patient/citizen, organization and economy) is to discover the "true cost" of health produced by application of that technology. After completing this module students and public health professionals from a variety of backgrounds should: Obtain an overview of a background and origins of health technology assessment, and an introduction to the scientific methods and instruments in health technology assessment; Summarize the four main elements of a health technology assessment analysis – the technology, the patient/citizen, the organization and the economy, and the steps that have to led to the assessed health technology; and Examine how decision-maker's questions are specified in health technology assessment, how literature is searched and collected, how studies could be designed, how data can be collected and analyzed, and how published health technology assessments could be validated.
Synopsis (Abstract)	Content of the course covers: Background, Origins, What is health technology?, What is health technology assessment?, What is the purpose of health technology assessment?, Is it health technology assessment or a different approach that is needed?, What are the main elements of HTA analysis?, When are health technologies assessments requested?, What is the role of ethics in health technology assessment?, How is health technology assessment conducted?, Selected issues in health technology assessment, Case example, and References.
Teaching methods	Lectures; Study of literature in small groups (up to five students); Guided discussion on previously done exercises and case problems; to prepare a project report (in a group of three students) on one topic for a certain health technology.
Specific recommendations for teacher	The topic allows a good combination of theoretical knowledge with practical skills. Knowledge in quantitative and qualitative research designs and measurement issues; various methods from statistical, over informatics to economic is already expected from the student, as well as skills in computer and language. This module should be only first in line with three lectures, six exercises and group/individual work (three times).
Assessment of Students	Project work with defense of the study; Multiple-choice questionnaire.

INTRODUCTION TO HEALTH TECHNOLOGY ASSESSMENT

Jelena Marinkovic

Worldwide, publicly funded health services are under pressure due to demographic changes, growing expectations, and the development of new technologies. Of these three major pressures, new technologies are generating the most concern and the most dramatic response (1-3). Still, new technologies can benefit health and disability service consumers in many ways. Some directly improve quality of life or life expectancy. Others act more indirectly, for example, by increasing the efficiency of the health system. However, new technologies are often introduced before there is adequate information about safety, effectiveness and ethical and social acceptability.

Origins

Technology assessment (TA) arose in the mid-1960s from an appreciation of the critical role of technology in modern society and its potential for unintended, and sometimes harmful, consequences. Experience with the side effects of a multitude of chemical, industrial and agricultural processes, and such services as transportation, health and resource management contributed to this understanding (4). TA was conceived as a way to identify the desirable first-order, intended effects of technologies as well as the higher-order, unintended social, economic and environmental effects (5).

Health Technologies (HT) had been studied for safety, effectiveness, cost, and other concerns long before the advent of Health Technology Assessment (HTA). Development of TA as a systematic inquiry in the 1960s and 1970s coincided with the introduction of health care technologies that prompted widespread public interest in matters that transcended their immediate health effects. Health care technologies were among the topics of early TAs. Multiphasic health screening, in vitro fertilization, predetermination of the sex of children, retardation of aging and modifying human behavior by neurosurgical, electrical or pharmaceutical means were among the first, "experimental", assessments.

Since its early years, HTA has been fueled in part by emergence and diffusion of technologies that have evoked social, ethical, legal, and political concerns. Among these technologies are contraceptives, organ transplantation, artificial organs, life-sustaining technologies for critically or terminally ill patients, and, more recently, genetic testing and genetic therapy. These technologies have challenged certain societal institutions, codes, and other norms regarding fundamental aspects of human life such as parenthood, heredity, birth, bodily sovereignty, freedom and control of human behavior, and death (6).

HTA is the only field of TA so far which has gained a distinctive profile in the sense of a particular subject, client, expertise and specialized institutions. HTA like TA in general aims at supporting decision making by providing comprehensive information on the preconditions for, and consequences of the implementation of new technologies (7).

What is health technology?

Goodman defines technology as the application of scientific or other organized knowledge--including any tool, technique, product, process, method, organization or system--to practical tasks. In health care, technology includes drugs; diagnostics, indicators and reagents; devices, equipment and supplies; medical and surgical procedures; support systems; and organizational and managerial systems used in prevention, screening, diagnosis, treatment and rehabilitation (4).

According to the World Health Organization's (WHO) "Health for all Policy in the 21st Century", released in January 1998, the scope of technologies for health, extends from those technologies that provide a direct benefit to health (such as molecular genetics, biological, pharmaceuticals, and medical devices), to those that support health system functions (like telecommunications, information technologies, devices for environmental protection and food technologies).

The International Network of Agencies for Health Technology Assessment (INAHTA) defines health technology as prevention and rehabilitation, vaccines, pharmaceuticals, and devices, medical and surgical procedures, and the systems within which health is protected and maintained (8).

Under broad definitions such as these, the phrase "health (healthcare, medical) technology" can be used in both diagnostic and therapeutic settings, and under either individual or population health approaches. Health technologies might include, for example, chemotherapy for cancer, hearing aid technology, electronic fetal monitoring, population screening for breast cancer, coronary artery bypass surgery, and magnetic resonance imaging.

As the field of health technology assessment has evolved, these definitions have come to be seen as a fairly narrow definition of technology. In part, this has been due to a growing recognition that the arrangements and structures for delivery of drugs, device and procedures can have far reaching impacts not only on the use of technology but also outcomes of patients. To reflect the importance of these and other

factors a more comprehensive definition of health technology is given by Kristensen. He defines very broadly that health technology is the practical application of knowledge in relation to health and disease (9).

With the health problem as the starting point, according to Bakketeig (10), the aim of the technology can roughly be divided into following: preventive care (aimed at preventing diseases from occurring), screening (aimed at detecting early signs of diseases or risk factor, with the aim to slow down the development of the disease), diagnosis (aimed at identifying the diseases in patients with clinical signs and symptoms), treatment (seeking to maintain health status, cure the patient or provide palliation), rehabilitation (which takes its starting point in the treated, but still ill patient and seeks to restore the functioning or minimize the consequences of dysfunction or defects).

What is health technology assessment?

While there is no widespread consensus on the definition of health technology assessment, for a long time a widely accepted definition was that of the United States Office of Technology Assessment (OTA) that it is the field of research that evaluates the short and long-term consequences of individual medical technologies on individuals and society (11).

HTA is related to research due to its methods, but is also related to planning, administration, and management due to its focus on decision-making. Thus, HTA can be seen as a bridge between science paradigm and a policy paradigm (12).

In Europe in mid nineties HTA is seen as a structured analysis of a health technology, a set of related technologies, or a technology-related issue that is performed for the purpose of providing input to a policy decision (13). HTA beside the benefits and financial costs of a particular technology or group of technologies also includes studies of ethical and social consequences of technology; factors speeding or impeding development and diffusion of health technology; the effects of public policies on diffusion and use of health technology and suggested changes in those policies; and studies on variation in use of technologies (13).

Goodman defines HTA as a systematic evaluation of properties, effects, and/or impacts of health technology. It may address the direct, intended consequences of technologies as well as their indirect, unintended consequences. Its main purpose is to inform technology-related policymaking in health care. HTA is conducted by interdisciplinary groups using explicit analytical frameworks drawing from a variety of methods (4).

The International Network on Agencies for Health Technology Assessment defines HTA as multidisciplinary field of policy analysis that studies the medical, social, ethical, and economic implications of development, diffusion, and use of health technology (8).

The broadest one is given by Kristensen who defines health technology assessment as a research based, applied assessment of relevant available knowledge of problems, when applying technology in relation to health and disease. HTA is a comprehensive, systematic assessment of the conditions for, and the consequences of using health technology (9).

What is the purpose of health technology assessment?

The purpose of HTA is to assist health policy makers, managers and health professionals at local and national levels in making informed decisions both in health purchasing, policy and practice. HTA information may be particularly useful in supporting decisions when: an established technology is associated with significant variations in utilization or outcomes, a technology is highly complex or involves significant uncertainty, a technology has high unit or aggregate cost, explicit trade-off decisions must be made in allocating resources among technologies, or a proposed provision is innovative or controversial.

The essential properties of HTA are the orientation to decision-making and its multidisciplinary and comprehensive nature. The goal of HTA is change. That is, it encompasses all methods used by health professionals to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

Is it health technology assessment or a different approach that is needed?

It is useful to clarify whether HTA is the right instrument to use for the particular problem because it may conceivably be more beneficial to apply a different approach. According to Kristensen et al. (9), the set of alternative procedures to clarify the problem are:

Alternative procedures to clarify the problem
1. Health technology assessment
2. A quality-assurance project (if one knows what should be done in particular organizational situation, but what is presently done, is not the right approach).
3. A basis for decision-making developed in the usual administrative framework (if, for instance, a national HTA or an HTA from region is available).
4. A traditional expert and/or stakeholder committee (if the aspect of stakeholders is very important, or if the opinion of particular expertise desired, or if only little time is available).
5. Exclusively a systematic literature review, possibly a meta analysis, to determine the clinical effects and efficiency of the technology.
6. An economic analysis (if sufficient knowledge of the effect and efficiency of the technology is available, and if there are no specific organizational questions).
7. A (primary-) research project (if documented research is simply not available, especially of the clinical effects).

Source: Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment, 2001.

What are the main elements of HTA analysis?

HTA includes analysis and assessment of a number of areas, where use of the health technology may have consequences. These can be divided into four main elements: the technology, the patient/citizen, the organization and the economy.

Technology. Assessment of technology includes the following main aspects that need to be assessed: field of application, effectiveness and risk assessment. This aspect is covered in more details later on.

Patient/Citizen. Examining the patient/citizen element in HTA is covered with very different methods, as regards theoretical basis and application; from field research that includes participant observation, interviews that include focus group, questionnaire surveys to prospective methods. Frequently measurement process is based on health status and health-related quality of life concept (4,14,15).

Organization. The aim of organizational analysis is to pinpoint some of the dimensions, which can be of importance for how interaction between technology, organization and administration develops, i.e. to describe some of the elements, which could play a part in the interaction between the behavioral patterns around technology, and point out possible consequences of different directions (16,17).

Economy. Economy aspect includes economic, budget or business analysis. The first one is far more important and is mainly conducted at societal level, where economic consequences for society, which means everyone who is directly or indirectly affected by technology, are assessed and included. Budget analysis is applied when investigating who carries the burden in terms of expenditures and who will benefit from the use of technology. At last, business analysis is conducted when the information about needs for investment and the running costs with respect to a technology are important (18).

When are health technology assessment requested?

Assessments can be requested and conducted at any stage in a technology's life cycle.
Stages include: conceptual (in the earliest stages of development), experimental or investigational (undergoing initial testing and evaluation), pre-established (adoption of an innovation by certain individuals and institutions); established (considered to be a standard approach and diffused into general use); and outmoded (superseded by another technology or demonstrated to be ineffective or harmful).

Since technology is constantly evolving, HTA must be viewed as an iterative process. It may be necessary to reassess a technology when competing technologies are developed, the technology itself evolves or new information is introduced.

What is the role of ethics in health technology assessment?

Since HTA is used to make judgments about what ought to be done with health technologies, there is significant overlap between it and medical ethics. According to Goodman (19) conducting a technology assessment requires careful attention to ethical questions, such as:

- Should all assessments be driven by cost concerns?
- Are the individuals involved in the selection of topics, the conduct of the assessment, and the use of its results free of conflicts of interest?
- Are judgments of value implicit in the statement of the assessment problem or the choice of methodology?
- Are informed consent, patient confidentiality and related means for protecting patient welfare in clinical investigations properly implemented?
- Do assessments provide means (e.g., in data collection, synthesis and reporting) to determine how technologies challenge prevailing legal standards and societal norms?
- Are the assessment's recommendations ethically justified?

How is health technology assessment conducted?

HTA process involves: the identification of technology, health or health care problems and possible assessments to address these; the priorization of possible assessments; assessment; dissemination of the

findings and conclusions of assessments; the implementation of findings and conclusions in policy and practice and impact assessment of resulting change.

The ten steps listed below, according to Goodman, provide a basic classical framework for conducting a health technology assessment (not all assessments involve each of these steps or conduct them in the same sequence) (19).

Ten Steps of Health Technology Assessment
1. Identify and rank assessment topics
2. Specify assessment problem
3. Determine locus or responsibility of assessment
4. Retrieve available evidence
5. Collect primary data (as appropriate)
6. Interpret evidence
7. Synthesize evidence
8. Formulate findings and recommendations
9. Disseminate findings and recommendations
10. Monitor impact of assessment reports

Source: Goodman C, Snider G, Flynn K. Health Care Technology Assessment in CVA. Boston, Mass: Management Decision and Research Center, Washington, DC: Health Services Research and Development Service, 1996.

When new technology is in question this framework has somewhat different stages:

A Framework for New Technology Assessment
1. Horizon scanning - the identification of emerging technologies before they become available for introduction.
2. Prioritization for assessment - deciding which new or emerging technologies should undergo further assessment.
3. Assessment - a research-based process designed to determine whether a new technology is safe, efficacious, effective and efficient.
4. Appraisal - a judgment on the social and ethical acceptability and appropriateness of a new technology. This includes consideration of community need, equity, and opportunity cost.
5. Adoption and diffusion - the process whereby new technologies are taken up in clinical practice.
6. Evaluation - the ongoing assessment of a new technology following its introduction.

Source: National Health Committee. New Technology Assessment in New Zealand. Discussion document, 2002.

The source of next definitions and descriptions is publication “New Technology Assessment in New Zealand”, published in year 2002 (20).

Horizon scanning is the process of identifying new and emerging technologies that have the potential to impact on a health system. It essentially involves formal or informal communication between policy makers and experts (21).

New technology assessment requires significant time, expertise and resources. Consequently, it is impossible to assess all emerging technologies. Therefore, prioritization for assessment is an essential and crucial part of the framework. There should be an agreed set of criteria against which emerging technologies are prioritized for assessment.

The assessment stage is based on empirical research. It aims to establish the effect (safety, efficacy and effectiveness) and efficiency of a new technology. This phase can be costly and labor intensive as it may involve primary research. Where possible, a systematic review of the scientific evidence is performed (mainly using scientific literature from peer-reviewed journals).

The gold standard test for safety, efficacy and effectiveness is a double-blind randomized controlled trial (RCT). However, for several logistical and ethical reasons, it is not possible to conduct double-blind RCTs for all new technologies. Where double-blind RCTs cannot be carried out, it is necessary to rely on the best alternative source of evidence. The task of selecting the best source of evidence is made easier by using well accepted, levels of evidence (22).

The efficiency (value for money) of a new technology is predicted by economic evaluation. There are several types of economic analysis that can be used to determine “value for money”. These include Cost-Minimization Analysis (CMA), Cost-Effectiveness Analysis (CEA), Cost-Benefit Analysis (CBA) and Cost-Utility Analysis (CUA). These may be referred to collectively as efficiency analysis.

The appraisal of a new technology involves taking into consideration community need, equity, appropriateness and acceptability, and opportunity cost. In contrast to the assessment phase, appraisal is more of an art than a science. It requires judgments to be made on social values and is informed by understanding of the health and disability sector and society in general. Inputs from professionals, consumers and the wider community are considered to be particularly important at the appraisal stage.

The adoption and diffusion stage is relatively self-explanatory. As a new technology appears to be of value, patients begin to request it and clinicians begin to use it. Ideally, a new technology that has been

assessed and appraised and found to have a potential benefit will be adopted and diffused into the health and disability sector in a controlled manner; that is, the circumstances in which it is used will be agreed on before the technology has been adopted and diffused. In reality, new technologies tend to be adopted and diffused in a rather disorganized manner.

New technology assessment should be an iterative process rather than a one-off study. Evaluation helps to ensure that this is the case. It involves the monitoring and further studying of a technology once it has been introduced. This might include: preparation of qualitative and quantitative data collection systems to receive data for side effects and complications, appropriateness and acceptability of the community for the new technology and outcome measures conducting scheduled milestone evaluation to determine achievement of target evaluation measures collaboration with clinical evaluation and quality improvement programs (23). Evaluation is important because the disease patterns and other characteristics of the population using the technology will inevitably change, and this may have implications for safety, effectiveness, efficiency and so on. In addition, during the initial stages of a technology's life cycle, the skills of practitioners in using the technology are not likely to be much higher. As skill level increases, the balance of risks and benefits associated with the new technology may change considerably.

Selected issues in health technology assessment

Quality of Care and HTA. Quality of care is a measure or indicator of the degree to which health care is expected to increase the likelihood of desired health outcomes and is consistent with standards of health care. Quality assurance involves a measurement and monitoring function, (i.e., quality assessment). HTA and quality assurance are distinct yet interdependent processes that contribute to quality of care. HTA generates findings that add to our knowledge about the relationship between health care interventions and health care outcomes. This knowledge can be used to develop and revise health care standards, e.g., manufacturing standards, clinical laboratory standards, practice guidelines, and other agreed upon criteria, practices and policies regarding the performance of health care. In summary, HTA contributes knowledge used to set standards for health care, and quality assurance is used to determine the extent to which health care providers adhere to these standards (24,25).

Outcomes Research and HTA. Outcomes research concerns any inquiry into the health benefits of using a technology for a particular problem under general or routine conditions (26). In practice, the term outcomes research has been used interchangeably with the term effectiveness research since the late 1980s to refer to a constellation of methods and characteristics that overlap considerably with HTA.

Centralization and decentralization of HTA. Although technology assessment have originated as a centralized function conducted by government agencies or other national- or regional-level organizations, HTA is also a decentralized activity conducted by a great variety of organizations that make technology-related policy decisions (27). As noted before, a HTA done from a particular perspective may not serve the technology-related policymaking needs of other perspectives.

Evidence-based health technology assessment. Eisenberg considers the next ten lessons for evidence-based technology assessment: innovation and flexibility should guide assessment; technology is more than devices; research and assessments should be linked with coverage; technology assessment is not a one-time exercise; new measures of outcomes should be developed; the community of practice is a laboratory for technology assessment; training and capacity building in technology assessment should be emphasized; better international collaboration will result in global synergy; national resources on technology assessment should be linked and technology assessments should be translated into improved practice (28). The same author writes, "Evidence-based technology assessment is a critical public good that can benefit all who are concerned about appropriate use of health services and products. Technology itself is rarely inherently good or bad, always or never useful. The challenge is to evaluate when it is effective, for whom it will enhance outcomes, and how it should be implemented or interpreted. Health technologies will not reach their potential unless they are translated, used, and continuously evaluated".

Institutionalization of Health Technology Assessment

The growth of HTA internationally can be seen in the growth of the International Network of Agencies for Health Technology Assessment (INAHTA; http://www.inahta.org), established in 1993. Organizations and individuals involved in HTA research are also affiliated with societies such as the Health Technology Assessment International (HTAi; http://www.htai.org), the International Society for Pharmacoeconomics and Outcomes Research (ISPOR; http://www.ispor.org), WHO Collaborating Center for Knowledge Translation and Health Technology Assessment in Health Equity (http://www.cgh.uottawa.ca/whocc/), or the European network for health technology assessment (EUnetHTA; http://www.eunethta.eu).

The governmental organizations worldwide that include topics on HTA are: Agency for Healthcare Research and Quality (AHRQ; http://www.ahrq.gov/research/findings/ta/index.html), Canadian Coordinating Office for Health Technology (CADTH; http://www.cadth.ca/en/products/health-technology-

assessment), Centre for reviews and Dissemeniation at University of York (CDR; http://www.york.ac.uk/inst/crd/), UK National Institute for Health and Care Excellence (NICE); http://www.nice.org.uk/guidance/index.jsp?action=byType&type=6), and UK National Institute for Health Research (NIHR; http://www.hta.ac.uk).

The main information on HTA can be found at (a) The University of York Centre for Reviews and Disseminations (CRD) HTA database which has over 12,000 summaries of completed and ongoing health technology assessments that can be retrieved from the http://www.york.ac.uk/inst/crd/, or appropriate journals (b): Health Technology Assessment, The internationally acclaimed journal series of the NIHR HTA Programme (http://www.hta.ac.uk/research/HTAjournal.shtml), International Journal of Technology Assessment in Health Care (journals.cambridge.org/action/displayJournal?jid=THC), and GMS Health Technology Assessment (http://www.egms.de/en/journals/hta/).

An international masters program in health technology assessment and management, ULYSSES, is also offered (http://www.ulyssesprogram.net), as well as HTA 101 and Comparative Effectiveness Research a Two-Part Webinar Series (http://www.nlm.nih.gov/nichsr/htawebinars/).

Case Example

The example is related to computer-based delivery of health evidence done as a health technology assessment in report "Computer-Based Delivery of Health Evidence: A Systematic Review of Randomized Controlled Trials and Systematic Reviews of the Effectiveness on the Process of Care and Patient Outcomes" done by Cramer et al from The Alberta Research Centre for Child Health Evidence, University of British Columbia, 2003, (29). The basic framework and the explanation are suggested by Goodman, (19).

Step 1. Identify and rank assessment topics.

Identifying potential topics. /To a large extent, assessment topics are determined, or at least bounded, by the mission or purpose of an organization. /

The perspectives opened up by information and communication technology for health and health care go beyond problems of the clinical setting and relate health to general problems of the so-called Information Society. Over the past decade, in an effort to assist health professionals with successfully searching for, translating, and integrating the best clinical evidence at the point-of-care, computer-based evidence delivery systems have been developed. These systems have been designed to assist providers with diagnosis, prescription, managing diseases, and preventing diseases. In addition to assisting health professionals, these systems have been designed to assist health care consumers by guiding them in their health behaviors, treatment options and disease management.

Ranking topics. /Some assessment programs have explicit procedures for setting priorities. Others set priorities in ad hoc or informal ways. The following are examples of criteria — listed in no particular order — that might be used to set assessment priorities: high burden of morbidity or mortality; large number of patients affected; high unit or aggregate cost of a technology or health problem; substantial variations in practice; high potential to improve health outcomes or reduce health risks; availability of sufficient research findings to perform the assessment; scientific, professional or public controversy; need to make regulatory decision; need to make payment decision; available findings not widely disseminated or used by practitioners./

Selected topic fulfill most of the criteria listed above and has almost the greatest importance of all ICT application in the filed of health.

Step 2. Specify assessment problem. /One of the most important aspects of an assessment is to specify clearly the question(s) to be addressed; this will affect all subsequent aspects of the assessment. Assessment problem statements should recognize the relation of the new technology to existing technology./

As with any innovative health care intervention, computer-based evidence delivery system need to be rigorously evaluated before their use become widespread (get acquainted with a Health on the Net Foundation- HON – principles, http://www.hon.ch). The objective of this assessment was to systematically identify and synthesize randomized controlled trials (RCT) and systematic reviews (SR) that evaluate the effectiveness of computer- based health evidence delivery systems on the process of care (e. g., compliance with evidence) and/ or patient outcomes (e. g., blood pressure).

Step 3. Determine locus of assessment. /The nature of an assessment problem will affect the determination of the most appropriate organization or group to conduct the assessment. A comprehensive assessment addressing multiple attributes of a technology can be very resource-intensive. It can require considerable training and experience in the methods of evidence-based medicine. Factors that influence a HTA "make or buy" decision include: Is an existing assessment available? If an existing assessment is available, does it address the specific issues of concern to the organization? How recently was it conducted? Is the methodology used sufficiently credible? If an existing assessment needs to be updated or is not available, do people in the organization have the time and expertise to perform the required data collection and analyses? If a synthesis of existing information is needed, does the organization have database searching

capabilities, staff to retrieve full text articles, and staff trained in the conduct of systematic reviews? If new data are needed, does the organization have the requisite resources and expertise? What methodology will be used? If a consensus of clinical experts is the preferred methodology, does that consensus need to incorporate and reflect the opinions of the organization's own clinicians? Will local clinicians accept the results and report recommendations if they do not participate in the assessment?/

Step 4. Retrieve available evidence. /One of the great challenges in HTA is to assemble all of the evidence relevant to a particular technology before conducting a qualitative or quantitative synthesis. Although some sources are devoted exclusively to health care topics, others cover the sciences more broadly. Multiple sources should be searched to increase the likelihood of retrieving all relevant reports. Useful sources for relevant evidence include: computer databases of published literature; computer databases of clinical and administrative data; printed indexes and directories; government reports and monographs; reference lists in available studies, reviews and meta-analyses; special inventories of reports; health newsletters and newspapers; company reports; and colleagues and other investigators. Increasingly, most of the sources are accessible via the Internet./

Evidences are taken from published and unpublished randomized clinical trials and systematic reviews that assess the effectiveness of computer-based evidence delivery systems. In this reviews, a comprehensive search of the literature using following databases: Medline (1990-2002), EMBASE (1990-2002), CINAHL (1990-2002), Cochrane Controlled Trials Register (1990-2002), Web of Science (1990-2002), and the trial registry of the Cochrane Effective Practice and Organization of Care Group (1990-2002) was done. In addition, two reviewers independently hand-searched the Health Information and Libraries Journal (1990-2002), Journal of the Medical Library Association (1990-2002), Medical Reference Services Quarterly (1990-2002), and the Proceedings of the American Medical Informatics Association (1991-2002). In addition, individuals from companies (more than 60) that produce relevant products were contacted for information about relevant studies. Finally, authors of all relevant articles and experts in the field are being contacted for information on recent, ongoing, or unpublished studies. This comprehensive search of literature at last identified 13 570 documents of which 525 were deemed potentially relevant for the selected assessment question.

Step 5. Collect primary data. /Compiling evidence for an assessment may entail collecting new primary data after determining that existing evidence will not adequately address the assessment question(s). Methods for generating new data on the effects of health technology ranges from case reports to meta-analysis. The demand for studies of higher methodological rigor (e.g., meta-analysis or RCTs) is increasing among health care technology regulators, payers, providers and other decision makers./

Step 6. Interpret evidence. /Evidence interpretation involves classifying the studies, grading the evidence and determining which studies will be included in the synthesis. Assessors should use a systematic approach to critically appraise the quality of the available studies. Interpreting evidence requires knowledge of investigative methods and statistics./

Two reviewers independently screened 525 articles for relevance using a predetermined set of inclusion criteria and identified 57 relevant randomized controlled trials (RCT) and 10 relevant systematic reviews. The majority of these studies was rated as having low methodological quality and was therefore open to substantial bias. The majority of the RCTs, as well as systematic reviews, were published between 1995-2001, (33 and 9 respectively), and were conducted in North America (46 and 6).

Step 7. Synthesize and consolidate evidence. /For many topics in technology assessment, a definitive study that indicates one technology is better than another does not exist. Even where definitive studies do exist, findings from a number of studies often must be combined, synthesized or considered in broader social and economic contexts in order to respond to the particular assessment questions. Methods used to combine or synthesize findings from different studies include: systematic reviews, meta-analysis, decision analysis and group judgment or consensus development./

One method for providing an evaluation is to summarize the existing evidence in a systematic review. Systematic reviews use explicit and reproducible methods for identifying and selecting primary or integrated studies and assess the methodological quality of each study with respect to the strength of evidence it contains.

Eighteen of the 57 randomized controlled trials investigated systems designed specifically for patient users, 37 studies investigated systems designed specifically for health care providers, and two studies investigated systems designed for use by both patients and health care providers. Five studies investigated diagnosis systems, 30 investigated management systems, one investigated a prediction system, four investigated prescription systems, nine investigated prevention systems, six investigated support systems, and two investigated treatment systems. The primary outcomes measured varied considerably from study to study and were categorized into one of three groups: process of care (e.g., compliance with medical guidelines), patient health (e.g., blood pressure), and other (e.g., knowledge).

When the data from these studies were pooled, use of these systems was found to enhance the process of care. However, some studies showed a positive effect of these systems on the process of care whereas other studies did not. The variability among the findings of these studies is likely a result of the various differences between the studies such as the intervention studied, the methodological quality, or the specific outcomes assessed. Overall, the use of computer-based evidence delivery systems was not found to have an impact on patient health outcomes. However, there were very few studies that investigated patient health outcomes and in most cases, the studies were too small to detect an effect. In addition, to have an effect on patient health outcomes, these systems must first have an effect on the process of care. Thus it may be too early to investigate patient health outcomes. The effect of these systems on the process of care needs to be enhanced prior to investigating their effect on patient health outcomes.

Six of the ten systematic reviews included studies with experimental designs other than randomized controlled trials and three of the ten assessed studies with designs other than controlled clinical trials. Two included investigations of non-computerized as well as computerized information systems. Eight reviews investigated the effects of these systems on the process of care and seven found a benefit. The effect of these systems on patient health outcomes was tested in eight systematic reviews and four documented a benefit. These findings are consistent with the findings of the review of randomized controlled trials.

Step 8. Formulate findings and recommendations. /Although the terms "findings" and "recommendations" are sometimes used interchangeably, they have different meanings. Findings are the results or conclusions of an assessment; recommendations are the suggestions, advice, or counsel that follow from the findings. Recommendations can be made in various forms, such as options, practice guidelines or directives./

Firstly, findings compromise that there exists great variability among these computer-based systems and the findings of the studies. Thus, there may not be one generic system that works in all environments. There is a need to identify factors that contribute to successful and unsuccessful systems. And, every system needs to be evaluated in the environment where it is implemented. Secondly, compliance with evidence is low with and without the use of these systems. Therefore, there is the need to identify barriers to the uptake of evidence, and where the barriers are inappropriate, to identify methods to remove them.

Broadly, several implications and recommendations for future areas of research can be suggested from this review. First, there is considerable potential for improving the dissemination and use of medical evidence. Future studies employing a qualitative approach are required to identify the barriers to using medical evidence and, where these barriers are inappropriate, the methods to remove them. In addition, because the results of the included studies varied (i.e., some found a benefit of using a computer-based evidence delivery system others did not) further research needs to focus on identifying the specific aspects of a system that contribute to its success or failure. This information will prove key to developing and implementing computer-based evidence delivery systems in the future.

Step 9. Disseminate the findings and recommendations. /Dissemination strategies depend upon the mission or purpose of the organization sponsoring the assessment. Dissemination should be planned at the outset of an assessment along with other assessment activities and should include a clear description of the target audience as well as appropriate mechanisms to reach them. The costs, time and other resources needed for dissemination should be budgeted accordingly. Dissemination plans do not have to be rigid. The nature of the findings and recommendations themselves may alter the choice of target groups and the types of messages to be delivered. Dissemination should be designed to influence the behavior of relevant decision makers./

New primary studies, new technology assessments, new policy on ICT by increasing relevance and delivery of information to health professionals and health consumers (get information on Health InterNetwork – HIN – United Nations Millennium Action Plan, http://www.healthinternetwork.org/index.php).

Step 10. Monitor impact of assessment reports. /The impact of HTAs is variable and inconsistently evaluated. Plans for monitoring the impact of an assessment report should be considered in the assessment design. Some of the effects of a HTA report include: acquisition or adoption of a new technology; reduction or discontinuation in the use of a technology; change in behavior; change in the organization or delivery of care; reallocation of national or regional health resources; change in regulatory policy; modification of marketing plan for a technology, … ./. Yet too early to say in this case example.

Exercises

A. Selection and prioritization

1. Identify possible health technologies in your country, region or institution that would be worth of assessment.

2. Define the criteria for prioritization and select the one technology worth assessing.

B. Planning/policy questions

1. Should there be a wish to introduce a public offer of influenza vaccination of the elderly, how should this be organized and what would the effects and costs be?

C. HTA questions

1. Derive HTA questions for influenza vaccination of the elderly keeping in mind that they have to be clearly worded, defined, answerable and limited in number.

D. Define Project group

1. Define complete project management for assessment of influenza vaccination of the elderly.

E. A HTA is to a large extent based on available evidence.

1. List possible sources for any literature review.
2. Perform a literature review with previously defined search protocol for "Influenza vaccination of the elderly" concerning HTA question – technology: What is the expected survival of the elderly, who are vaccinated against influenza, compared to elderly, who are not vaccinated?
3. Perform a literature review with previously defined search protocol for "Influenza vaccination of the elderly" concerning HTA question – patient: What do the elderly think of influenza vaccination?
4. Define and perform a literature review with the previously defined search protocol for "Influenza vaccination of the elderly" concerning HTA.
5. Define and perform a literature review with the previously defined search protocol for "Influenza vaccination of the elderly" concerning HTA.

F. If the literature review didn't give enough scientific documentation there is a need for performing one's own study of the effect of health technology.

1. Design studies for HTA questions cited in E2-E5.
2. What are possible sources of bias in selected designs?
3. Define advantages and disadvantages in selected designs.
4. How would you measure validity in previously designed studies?
5. If you choose to measure health status what type of instruments can you use?

Case problems:

A. Most of the studies on health technology assessment covered new therapeutic and diagnostic health technologies and medical treatments basically concerning economic aspect and medical or patients benefits. This is a common result of few studies realized and published in mid nineties. What is situation today?

B. The inclusion of an unbiased sample of relevant studies is central to the validity of systematic reviews and meta-analysis. Time-consuming and costly literature searches, which cover the grey literature and all relevant languages and databases, are normally recommended to prevent reporting biases. However, the size and direction of these effects is unclear at present. There may be trade-offs between timeliness, cost and the quality of systematic reviews. It seems that there has to be an answer on the question: How important are comprehensive literature searches and the assessment of trial quality in systematic reviews?

C. Telehealth has become widespread in the last two decades in developed countries, despite the generally poor scientific evidence available to support its use. Telehealth, telemedicine, or e-health is defined as the use of information and communication technologies to deliver health services, expertise and information over distance, geographic, time, social and cultural barriers. Telehealth encompasses Internet or web-based "e-health", as well as video-based applications. Applications can be real-time or store-and-forward. How would you provide an information base to assist policy- and decision-makers, researchers and health professionals in their deliberation about telehealth? Provide an overview of the areas of strength and weakness, identify gaps and review policy implications.

D. Screening for gestational diabetes mellitus has been controversial, with some expert bodies advising universal screening, others selective screening, and yet others advising against screening at all. This has partly been a result of debate about the definition of gestational diabetes mellitus, and partly because of the profusion of different tests available, both for screening and definitive diagnosis. In the country X, there is no national policy on screening, and a variety of practices exist in different parts of the country. There have also been doubts about the treatment of gestational diabetes mellitus, and particularly about management of minor degrees of glucose elevation, which are better described as glucose intolerance rather than true diabetes. Provide an updated review of current knowledge, to clarify research needs, and to assist with policy making.

References

1. Perry S, Thamer M. Medical innovation and the critical role of health technology assessment. JAMA 1999, 282(19):1869-1872.
2. Fuchs VR, Garber AM. The new technology assessment. N Engl J Med 1990:323:673-677.

3. Perry S, Gardner E, Thamer M. Status of health technology assessment worldwide. International Journal of Technology Assessments in Health Care 1997:13:81-89.
4. Goodman C. TA101. Introduction to health care technology assessment. National Information Center on Health Services Research & Health Care Technology (NICHR),), 2004. National Library of Medicine: http://www.nlm.nih.gov/nichsr/hta101/ta101_c1.html.
5. Brooks H, Bowers R. The Assessment of Technology. Science 1970:222 (2):13-20.
6. National Research Council, Committee on the Life Sciences and Social Policy. Assessing Biomedical Technologies: An Inquiry into the Nature of the Process. Washington, DC: National Academy of Sciences, 1975.
7. Hennen L. TA in Biomedicine and healthcare – from clinical evaluation to policy consulting. TA-Datenbank-Nachrichten, Nr. 1/10: 13-22, 2001. Available at http://www.itas.fzk.de/deu/tadn/tadn011/henn01a.pdf.
8. The International Network of Agencies for Health Technology Assessment (INAHTA), http://www.inahta.org/.
9. Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment, 2001.
10. Bakketeig L. The Technology. In: Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment; 2001. p. 22-35.
11. Donaldson MS, Sox HC. Setting priorities for health technology assessment: A model process. National Academy Press, Washington, D.C. 1992.
12. Battista RN, Hodge MJ. The development of the Health Care Technology Assessment: An International Perspective. International Journal of Technology Assessment in Health Care 1995:11(2):287-300.
13. EUR ASSESS Project Subgroup. Introduction to the EUR ASSESS Report. International Journal of Technology Assessment in Health Care 1997:13:133-143.
14. Poulsen B. The Patient: Measurement of Health Status. In: Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment, 2001. p. 56-65.
15. Timm H, Hanses HP, Morgall J, Sigmund H. The Patient: Field Research, Interview and Questionnaire Studies. In: Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment; 2001. p. 37-55.
16. Vrangboek K. The Organization: HTA Administration and Organization. In: Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment; 2001. p. 67-84.
17. Tryggestad K, Borum F. The Organization. In: Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment; 2001. p. 85-95.
18. Poulsen PB. The Economy. In: Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment; 2001. p. 96-121.
19. Goodman C, Snider G, Flynn K. Health Care Technology Assessment in CVA. Boston, Mass: Management Decision and Research Center, Washington, DC: Health Services Research and Development Service, 1996.
20. National Health Committee. New Technology Assessment in New Zealand. Discussion document, 2002.
21. Banta D, Gelijns A. An early system for the identification and assessment of future health care technology: The Dutch STG Project. International Journal of Technology Assessment in Health Care 1998:14(4):607-612.
22. Centre for Evidence-Based Medicine. Levels of Evidence and Grades of Recommendations, 2011, http://www.cebm.net/index.aspx?o=1025 and 5513.
23. Stevens A, Robert G, Gabbay J. Identifying new health care technologies in the United Kingdom. International Journal of Technology Assessment in Health Care 1997:13(1):56-67.
24. Lohr KN ed. Institute of medicine. Medicare: A Strategy for Quality Assurance. Volume I. Washington, DC: National Academy Press, 1990.
25. Lohr KN, Rettig RA, eds. Quality of Care and Technology Assessment. Report of a Forum of the Council on Health care Technology. Washington, DC: National Academy Press, 1988.

26. Brook RH, Lohr KN. Efficacy, Effectiveness, Variations, and Quality: Boundary-crossing Research. Medical Care 1985:23(5):710-722.
27. Rettig RA. Healthcare in Transition: Technology Assessment in the Private Sector. Santa Monica, Ca.: RAND, 1997.
28. Eisenberg JM. Ten Lessons for Evidence-based Technology Assessment. JAMA 1999:282(19):1865-1869.
29. Cramer K, Hartling L, Wiebe N, Russell K, Crumley E, Pusic M, Klassen TP. Computer-Based Delivery of Health Evidence: A Systematic Review of Randomised Controlled Trials and Systematic Reviews of the Effectiveness on the Process of Care and Patient Outcomes. Alberta Heritage Foundation for Medical Research, 2003. Available at http://www.assembly.ab.ca/lao/library/egovdocs/alhfm/2003/146645.pdf.

Recommended Readings

1. Goodman C. TA101. Introduction to health care technology assessment. National Information Center on Health Services Research & Health Care Technology (NICHR), 2004. National Library of Medicine: http://www.nlm.nih.gov/nichsr/hta101/ta101_c1.html.
2. Kristensen FB, Horder M, Poulsen PB (eds.). Health Technology Assessment Handbook. Danish Institute for Health Technology Assessment, 2001.
3. Sorenson C, Drummold M, Kanavos P. Ensuring Value for Money in Health Care: The Role of Health. Technology Assessment in the European Union. Euro Nonserial Publication, 2008. ISBN 978-9-28907183-3.
4. HTA Core Model for rapid relative effectiveness assessment of pharmaceuticals, 2013 (http://www.eunethta.eu/outputs/new-application-hta-core-model-hta-core-model-rapid-relative-effectiveness-assessment-pharma).
5. HTA Core Model for screening technologies 2012 (http://www.eunethta.eu/outputs/final-hta-core-model-screening-technologies).
6. HTA Core Model on Abdominal Aorta Aneurysm Screening (http://meka.thl.fi/htacore/106.aspx).
7. HTA Core Model for medical and surgical interventions (http://www.eunethta.eu/outputs/hta-core-model-medical-and-surgical-interventions-10r).
8. HTA Core Model for diagnostic technologies (http://www.eunethta.eu/outputs/hta-core-model-diagnostic-technologies-10r).

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**SWOT analysis**
Module: 2.35	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Andrej Plesnicar, MD, PhD, Teaching Assistant Faculty of Health Studies, University of Ljubljana, Slovenia Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Andrej Plesnicar Faculty of Health Studies, University of Ljubljana, Slovenia Zdravstvena pot 5, Ljubljana, SI-1000 Slovenia E-mail: andrej.plesnicar@vsz.uni-lj.si
Keywords	Strategic planning, SWOT analysis.
Learning objectives	After completing this module students should: • understand the role of strategic planning in public health; • be familiar with SWOT analysis as one of situation assessment tools in strategic planning; • be able to perform a SWOT analysis.
Abstract	SWOT analysis is a strategic planning tool used to evaluate different types of positive and negative factors which may affect desired future outcomes of a project, programme, process, etc. As a case study, a SWOT analysis of the 2nd PH-SEE Summer School for Public Health Professionals is presented. The study was performed by the Slovenian participants of the course at the end of the course.
Teaching methods	An introductory lecture gives the students a first insight into the characteristics of SWOT analysis. The theoretical knowledge is illustrated by a case study. In continuation, students first carefully read the recommended readings. Afterwards they discuss the characteristics of SWOT analysis with other students. Afterwards, in small groups (2-3 students) they perform a SWOT analysis of a given example.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room, rooms for small-group work; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	SWOT analysis of a situation of a selected process.

SWOT ANALYSIS

Andrej Plesnicar, Lijana Zaletel-Kragelj

Theoretical background

Basic considerations

SWOT analysis definition

SWOT analysis is a strategic planning tool used to evaluate different types of positive and negative factors which may affect desired future outcomes of a project, programme, process, etc. SWOT stands for (1-3):

- S – strengths,
- W – weaknesses,
- O – opportunities, and
- T – threats.

Before describing the characteristics of SWOT analysis, it is important to place this analysis in a wider frame of strategic planning.

Strategic planning definition and basic characteristics

There exist several more or less similar definitions of strategic planning, two of them being:

- according to Geyer (2), strategic planning is a process, by which an organization analyses whether it is effective in its goals and objectives, and it establishes whether the organization needs to change its direction to fulfill its purpose. Strategic planning helps to respond to the external environment in the most effective way.
- according to Lerner (3), strategic planning is a complex and ongoing process of organizational change. It is based on thorough analysis of foreseen or predicted trends and scenarios of the possible alternative futures, as well as the analysis of internal and external data. It is a qualitative, idea driven process. It integrates "soft" data, not always supported quantitatively, such as experiences, intuition, and ideas, involves the organization in the ongoing dialogue, and aims to provide a clear organizational vision and focus. It is an ongoing, continuous learning process, an organizational dialogue, which extends beyond attaining a set of predetermined goals. It aims to change the way an organization thinks and operates, and create a learning organization;

SWOT analysis in the frame of strategic planning

Strategic planning is a several steps process. According to Geyer (2), these steps are:

1. Getting organized
2. Analysis of the situation
3. Development of a strategy
4. Drafting a plan
5. Approval of the plan, and
6. Implementation of the plan

One of the most important steps is analysis of the situation (2). SWOT analysis is one of many tools which could be used in this step of strategic planning, but SWOT analysis is by far the most popular (3).

Description of SWOT analysis

Factors, assessed in SWOT analysis could be internal or external. This tool involves specifying the objective of the project and identifying four groups of factors:

- positive (favourable) internal factors or strengths,
- negative (unfavourable) internal factors or weaknesses,
- positive (favourable) external factors or opportunities, and
- negative (unfavourable) external factors or threats.

Figure 1. SWOT analysis matrix

	Positive factors	Negative factors
Internal factors	STRENGTHS	WEAKNESSES
External factors	OPPORTUNITIES	THREATS

These four groups of factors could be presented in a SWOT analysis matrix (Figure 1).

SWOT analysis has been widely employed in industry and management as a decision making aid and as an introduction to planning in various applications. It's been also used to uncover new outlooks and identify problems that would impede progress (1).

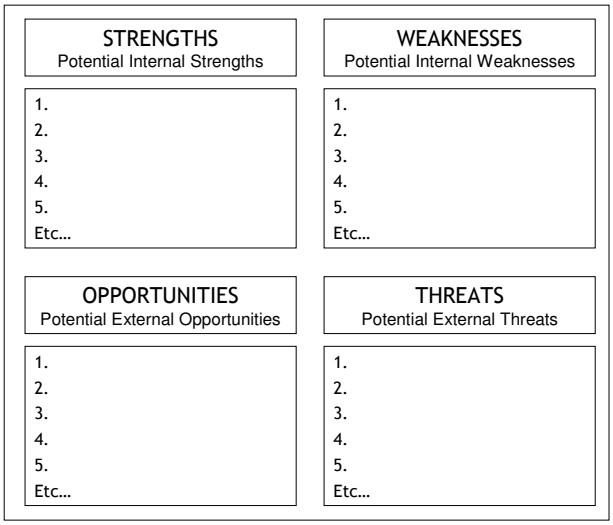
Doing SWOT analysis

SWOT analysis is a simple and quick method. The first step would be to make a simple worksheet by drawing a cross, creating four cells: one each for Strengths, Weaknesses, Opportunities and Threats. The next step is to list specific items related to the problems of the programme into the appropriate cells in the worksheet. Alternatively, a more detailed SWOT analysis can be made with the use of the template of a more elaborate worksheet scheme (Figure 2).

It is advised that no more than ten exact points should be listed under each heading (1); all of them should be carefully evaluated in the following sequence:

1. Do the internal analysis
   Internal analysis examines the advantages and drawbacks of the programme. This can be achieved with the analysis of strengths and weaknesses of the programme during its course and the impact they have on participants' perception.
2. Do the external analysis
   External analysis examines the main points in the analysis of the actual professional circumstances and environment of the participants. It also identifies those points that represent opportunities, and those that represent threats or obstacles to their future achievements. Thus, classify the answers or the data collected as external opportunities or threats.
3. Enter the collected information into the worksheet (Figure 2)
4. Apply the collected information to devise a strategy that uses the strengths and opportunities to reduce the *weaknesses and threats, and to achieve the objectives of the education programme in the participants' future* professional performance.

Figure 2. SWOT analysis worksheet template



SWOT analysis in public health

Strategic planning and particularly SWOT analysis could be used in public health, since public health problems require strategic planning. Among others, strategic planning could be useful methodology in health promotion, especially in mobilizing community partnerships to identify and solve health problems of a population, and informing, educating people about health issues. In fact, SWOT analysis is used in public health issues. One of examples is SWOT analysis of Halton's Healthy Living Project (4). However, SWOT analysis could be used in other fields of public health as well (5,6).

In short, an analysis of "internal" (Strengths and Weaknesses) and "external" (Opportunities and Threats) factors affecting any public health programme may help in making a plan for the future. Irrelevant and outdated programmes could thus be replaced with innovative and competent ones.

SWOT analysis could be also a method for assessing the quality and relevance of public health education programmes. It can be used by groups or individuals as an external assessment of such a programme, or may represent introspection into negative and positive concerns.

Case study

Swot analysis as an evaluation tool of public health education programmes - participants' assessment of 2nd ph-see summer school for public health professionals

Introduction

Parliamentary democracy and free market economy were reintroduced to the countries of South Eastern Europe (SEE) in the last decade of 20th Century after an interval of more than 50 years. In the majority of these countries democratically elected governments assumed responsibility in the late eighties and early nineties, and the region has been referred to as being in political and social transition. It is important to note that epidemiological transition had in some of the former socialist countries of SEE occurred long before the beginning of these changes. However, the problems associated with economic restructuring have been in these countries often compounded by the inability of inefficient governments and judiciaries to deal with them properly. Decreasing gross domestic product, extensive environmental damage, rising international debt, massive unemployment resulting from closures of large and ineffective socialist style companies, and growing corruption have been in some cases exacerbated by political instability and war. The extent of social and political reforms and the ability to create a functioning market economy are therefore not uniform in SEE and the population of the region frequently has to cope with the aftermath of economic deprivation, alienation and social inequality. Not least of all, the health expenditures remained on almost basic levels in some of these countries and have contributed to a decade of disturbing trends in public health indicators. As a consequence, the awareness of the need to tackle these problems has gradually emerged among the health professionals and people of SEE itself and Europe as a whole (7,8).

The concept of Public Health has simultaneously evolved into a broader philosophy that is concerned not only with the provision of preventive and therapeutic health services, but also with other components important for the functioning of health care systems. These include the dilemmas associated with medical personnel and facilities, environment, life styles, economic support and just distribution of health services. The philosophy of New Public Health therefore links the traditional Public Health issues of epidemiology and hygiene with more recent issues of health promotion, social responsibility, economics and management (9). As the New Public Health is concerned with the totality of the health care systems, it corresponds well to the present day circumstances and health problems in SEE. It is thus imperative for the Public Health experts in SEE to continuously upgrade and broaden their knowledge, and to receive extensive parallel education in related fields of biomedical, social and anthropological sciences, economics, management, and advances in technology (5,10,11).

Recent developments in SEE countries have led the international community to establish a comprehensive, long-term conflict prevention strategy for this region. The activity of Stability Pact for SEE, a partnership of more than 40 countries and formed following the European Union initiative in Cologne on June 10th, 1999, has in this role replaced earlier haphazard reactive intervention policies in SEE. It supports the SEE countries in their "efforts to foster peace, democracy, respect for human rights and economic prosperity in order to achieve stability in the whole region" through the implementation and enhancement of regional co-operation, and serves as an instrument to co-ordinate and facilitate the projects of all its partners (12). With regard to the need of preparing public health professionals in SEE to upgrade their knowledge in order to follow and cope with the developments mentioned before, a PH-SEE (Public Health Collaboration in South Eastern Europe) initiative, including Programmes for Training and Research in Public Health, was developed within the framework of Stability Pact for SEE. The 2nd PH-SEE Summer School for Public Health Professionals (2nd PH-SEE-SSPHP) that took place on July 21st-27th, 2002, in Ljubljana, Slovenia, represents a part of the PH-SEE curriculum with clearly stated purposes and objectives (13). Accordingly, the participants were asked to collect and present their assessments and opinions of the 2nd PH-SEE-SSPHP immediately after the lectures and exercises ended.

Lecturers, participants and the set-up of the 2nd PH-SEE-SSPHP

The faculty of the 2nd PH-SEE-SSPHP included twelve invited lecturers. All of them were well known experts in Public Health and some of the related fields such as statistics, management, physics and information technologies in their home countries. Three of the lecturers came from Croatia, four from Germany, and five of them came from Slovenia. An additional lecture was given on an ad hoc basis by a participant from Macedonia. In addition, three members of the Organizing Committee were constantly available on the spot to help the participants with equipment or advice if needed.

The majority of the participants of the 2nd PH-SEE-SSPHP came from the SEE countries. The largest number was from Slovenia, while others came from Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Romania, Macedonia, Serbia (including participants from Voivodina) and Montenegro (Table 1). One of the participants came from Germany, and there were no participants from Republic of Moldova and Kosovo. Altogether, out of 32 participants 26 were females and six were males. All had university degrees in medicine and/or experience in dealing with issues of Public Health. Among the Slovenian participants, nine out of 14 worked at regional Institutes of Public Health. Only five out of 18 participants from other countries held a similar position, with the rest being employed in various government departments and universities of their respective countries. All participants were proficient in English.

Table 1. Regional distribution of the participants of the 2nd PH-SEE Summer School for Public Health Professionals, Ljubljana, Slovenia, July 21st-27th, 2002

Country	Number of participants
Albania	3
Bosnia and Herzegovina	1
Bulgaria	2
Croatia	1
Germany	1
Macedonia	1
Romania	2
Slovenia	14
Yugoslavia	7
Voivodina	3
Serbia	2
Montenegro	2

The 2nd PH-SEE-SSPHP took place in the premises of the National Health Insurance Institute (NHII) in Ljubljana, Slovenia. The venue of the 2nd PH-SEE-SSPHP was in the basement of the NHII building in the centre of the city in a well equipped and comfortable lecture room with effective air conditioning system. All the participants had an access to personal computers (PC), usually two per PC, and all the PCs were provided with relevant software programmes and internet links. The language of the 2nd PH-SEE-SSPHP was English, while some of the participants conversed among themselves in other languages as well. Catering was provided by the NHII for all of the participants.

Evaluation of Public Health education programmes and SWOT analysis

Involvement of students in medical education programmes planning and evaluation on all levels has by now become an accepted practice in some parts of the world. This type of evaluation of Public Health education programmes thus forms an integral part of the programmes' continuous professional and scientific quality assessment and assurance (7,1). It should be as accurate and comprehensive as possible, and it is advisable a certain sequence of steps in the analysis and assessment of a programme is followed (Table 2). It should not be taken as a test resulting in either "pass or fail".

On the last day the lecturers and organizers asked the participants to evaluate the 2nd PH-SEE-SSPHP and present their assessments and opinions in the final discussion of the course. The participants were divided into two groups: the first was made from those from Slovenia, and the second from those from other countries. The participants were asked to discuss the advantages and drawbacks of the 2nd PH-SEE-SSPHP and for this purpose the Slovenian participants chose to use the SWOT analysis in their assessment, as mentioned previously in the text. The participants had 15 minutes to prepare their evaluation immediately after the last lecture at the 2nd PH-SEE-SSPHP ended.

Table 2. A sequence of suggested steps in the analysis of a Public Health education programme.

	Step	Description
1.	Assessment of the Public Health education programme; overall assessment:	• assessment of the plans of the programme. • assessment of the procedures of the programme:
		– planned goals. – exercise objectives
2.	Assessment of the "host" facilities of the Public Health education programme:	• faculty of the programme. • premises available for the programme:
		– provision of resources (equipment, etc…). – familiarity with the circumstances and conditions
3.	Assessment of the participants of the Public Health education programme:	• capability to carry out tasks as instructed (exercises, etc…). • adherence to the agreed standards

Results of the SWOT analysis of the 2nd PH-SEE-SSPHP

The results of the evaluation are necessarily fragmentary and incomplete despite the fact that SWOT analysis was used. The factors affecting the 2nd PH-SEE-SSPHP are listed in each cell of the SWOT analysis template from top to bottom, according to their importance as perceived by the Slovenian participants (Figure 3).

It is the Slovenian participants' impression that the "host" qualities of the 2nd PH-SEE-SSPHP easily translate into majority of its Strengths. Highly qualified lecturers, well organized seminars with relevant practical exercises, study and exercise materials of good quality, presentations of carefully selected skills and knowledge relevant to the interests of the participants all attest to the success of the faculty of the Summer School, its International Programme Committee and its Organizing Committee in their efforts to offer the participants relevant up-to-date information. For example, practical exercises enabled the participants to prepare presentations of comparisons of health indicators to place their national pattern in the context of other countries with similar economic and social circumstances. Well organized course of the 2nd PH-SEE-SSPHP with exact timing, well equipped premises and excellent catering made the attendance easier and more comfortable for all the participants and added to its success. Although the largest part of participants came from Slovenia, the participants' regional distribution was still regarded to be good. Finally, in Slovenian participants' view, all the participants were highly motivated to attend the 2nd PH-SEE-SSPHP.

Figure 3. Results of the SWOT analysis of the 2nd PH-SEE Summer School for Public Health Professionals, Ljubljana, Slovenia, July 21st-27th, 2002, performed by Slovenian participants

STRENGTHS Potential Internal Strengths	WEAKNESSES Potential Internal Weaknesses
1. Highly qualified lecturers 2. Well organised seminars and relevant practical exercises 3. Well organised course of the Summer School with exact timing 4. Highly motivated participants with good regional distribution 5. Study and exercise materials of good quality 6. Presentation of skills and knowledge relevant to the interests of the participants 7. Well equipped premises 8. Excellent catering	1. No theoretical study workshops 2. No natural light
OPPORTUNITIES Potential External Opportunities	**THREATS** Potential External Threats
1. Possible communication with internationally acknowledged experts 2. Possible communication with other participants 3. Exchange of experience 4. Training the participants for actively searching for funds for attending and organising conferences in the future	1. Lack of money

The absence of theoretical study workshops was regarded by some of the Slovenian participants as one of the internal weaknesses of the 2nd PH-SEE-SSPHP. The problem was not discussed extensively during the SWOT analysis and the missing theoretical workshops were not clearly defined. In one participant's opinion it was suggested that relevant thematic roundtables should be organized on similar meetings in the future. The lack of natural light was regarded only as a minor drawback of the 2nd PH-SEE-SSPHP.

External opportunities also seem to be closely associated to the "host" qualities of the 2nd PH-SEE-SSPHP. Formal and informal discussions between the experts of the faculty and participants led to the exchange of valuable experience during the course of the 2nd PH-SEE-SSPHP that should be further encouraged in the future. Training the participants for actively searching for funds for attending and organizing conferences in the future was regarded as important by the majority of Slovenian participants, but for many of them it seemed overly complicated with little chance of success and not relevant in the context of their professional duties and priorities.

Among the threats the Slovenian participants emphasized the "lack of money". Although none of them had any particular financial problems with regard to the participation at the 2nd PH-SEE-SSPHP, it was regarded as a sort of an "overall" threat to the feasibility of similar Public Health education meetings in the future.

Discussion

Evaluation of the 2nd PH-SEE-SSPHP with SWOT analysis, as performed by the Slovenian participants, stressed its significance in the context of present day economic, social and political circumstances in the SEE countries. It confirmed the importance of continuing education in Public Health in its broader context and provided the participants with some important new practical information and skills. It is clear from the SWOT analysis that the 2nd PH-SEE-SSPHP with its Strengths and Opportunities forms a valuable part of the PH-SEE initiative and its Programmes for Training and Research in Public Health.

It can be seen from the SWOT analysis of the 2nd PH-SEE-SSPHP that no major modifications are needed for the future meetings of similar nature. The 2nd PH-SEE-SSPHP helped to increase the sense of self-esteem of the participants, reinforced their knowledge, and has probably motivated them to better performance. Many of them may in the future assume responsibilities as health planners, policy analysts and administrators in Public Health departments of their governments, and as researchers and teachers in educational institutions, communities, organized medical care groups and voluntary health organizations (9,15). Their expertise may thus facilitate the course of anticipated reforms needed to upgrade quality in health care systems that must cope with changes in the health needs of societies in SEE countries (6-8,13).

SWOT analysis was effectively used in the evaluation of the 2nd PH-SEE-SSPHP. However, as the circumstances change with the passage of time due to national and regional influences, an updated analysis should be repeated frequently (16). Committees and other groups could use a SWOT analysis as a way of introspection into positive and negative concerns when deciding about new Public Health education programmes in SEE and elsewhere. SWOT analysis can help decision makers and individuals expand their vision and facilitate necessary adjustments. In any case, decision making should build on Strengths, reduce Weaknesses, seize Opportunities and react to Threats (17).

In conclusion, it is safe to say that PH-SEE initiative with its Programmes for Training and Research in Public Health programmes makes an important contribution in achieving self-sustaining regional stability in SEE. The 2nd PH-SEE-SSPHP represents one of the steps ahead to achieve such stability, with all of the participants having a clearer vision where the region, its public health experts and health care systems should be heading.

Exercises

Task 1

Carefully read the part on theoretical background of this module. Critically discuss the characteristics of SWOT analysis with your colleagues.

Task 2

Carefully read the following two documents:

- CINDI World Health Organization, Regional Office for Europe. A strategy to prevent chronic disease in Europe. A focus on public health action. The CINDI vision. Copenhagen: WHO, Regional Office for Europe; 2004. Available from: URL: www.euro.who.int/document/e83057.pdf. Accessed: May 30, 2013.
- Maucec Zakotnik J, Fras Z, Zaletel Kragelj L. The WHO Countrywide Integrated Non-communicable Diseases (CINDI) Programme in Slovenia. In: Donev D, Pavlekovic G, Zaletel-Kragelj L (editors). Health promotion and disease prevention. A handbook for teachers, researches, health professionals and decision makers. Lage: Hans Jacobs Publishing Company, 2007. p.204-219. http://www.snz.hr/ph-see/Documents/Publications/FPH-SEE_Book_on_HP&DP.pdf. Accessed: May 30, 2013.

Task 3

Make groups of three to four students and make SWOT analysis of CINDI programme in Slovenia.

Task 4

Discuss results of your SWOT analysis with results of other student groups.

References

1. Johnson G, Scholes K, Sexty RW. Exploring strategic management. Scarborough, Ontario: Prentice Hall; 1989.
2. Geyer Y. Strategic planning. Handbook series for community-based organizations. Pretoria: IDASA, 2006.
3. Lerner AL. A Strategic Planning Primer for Higher Education. Northridge: College of Business Administration and Economics, California State University; 1999.
4. Samuels T, Perry C. A SWOT analysis of Halton's Healthy Living Project: prospects for sustainability. Executive Summary. Chester: The University of Chester; 2005.
5. Miller MG. Environmental metabolomics: a SWOT analysis (strengths, weaknesses, opportunities, and threats). J Proteom Res. 2007;6:540-5.
6. Dageid W, Duckert F. The process of evaluating a capacity-building support initiative for HIV positive South Africans. Eval Program Plann. 2007;30:381-91.
7. Chenet L, McKee M, Fulop N, Bojan F, Brand H, Hort A, Kalbarczyk P. Changing life expectancy in Central Europe: is there a single reason? J Public Health Med. 1996;18:329-36.
8. Plesnicar AF. A glimpse on the cancer problem in Latin America from a Central European perspective: II. Challenge Newsletter. 2000:2:14-22.
9. Tulchinsky TH, Varavikova EA. The new public health. An introduction for the 21st century. San Diego, CA: Academic Press; 2000.

10. Kohler L. Public health renaissance and the role of schools of public health. Europ J Public Health. 1991;1:2-9.
11. Kahn K, Tolmann SM. Planning professional education at schools of public health. Am J Public Health. 1992;82:1653-7.
12. Stability Pact [homepage on the Internet]. Available from URL: http://www.stabilitypact.org. Accessed: May 30, 2013.
13. Laaser U, Bardehle D. Health monitoring and reporting systems. Proccedings of the 2nd PH-SEE Summer School for Public Health Professionals, Ljubljana, July 21st-27th. Ljubljana: University of Ljubljana, Faculty of Medicine; 2002.
14. World Federation for Medical Education. World summit of medical education. Edinburgh, London: World Federation for Medical Education; 1993.
15. Roemer M. Genuine professional doctor of public health the world needs. J Nurs Scholarship. 1999;31:43-4.
16. Glass NM: Pro-active management: How to improve your management performance? East Brunswick, NJ: Nichols Publishing; 1991.
17. Balamuralikrishna R, Dugger JC. Swot analysis: a management tool for initiating new programs in vocational schools. J Vocat Tech Training. 1995;12:108-15.

Recommended readings

1. Geyer Y. Strategic planning. Handbook series for community-based organizations. Pretoria: IDASA; 2006.
2. Lerner AL. A Strategic planning primer for higher education. Northridge: College of Business Administration and Economics, California State University; 1999.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Community health - Public health research methods and practice**
Module: 2.36	ECTS (suggested): 0.3
Authors	Selma Sogoric, MD, MPH, PhD, Associate Professor "Andrija Stampar" School of Public Health, Medical School, University of Zagreb Aleksandar Dzakula MD, PhD, Teaching Assistant "Andrija Stampar" School of Public Health, Medical School, University of Zagreb
Address for Correspondence	Selma Sogoric "Andrija Stampar" School of Public Health, Medical School, University of Zagreb Rockefellerova 4, Zagreb, Croatia E-mail: ssogoric@snz.hr
Keywords	Community, equity, health promotion, methods, public health, society.
Learning objectives	After completing this module students and public health professionals should: • understand the meaning of community structure and dynamics for public health practice • increase knowledge on community structures, dynamics and research models in public health • recognize basic approaches for community based (and participatory) public health intervention • differentiate various approaches for applied community studies • identified relations between community research and practice • improve capability to work with and inside the community
Abstract	For few decades, the value of a community, empowerment, community-based care, population-based needs assessment was discussed, but not so much of the evidence of this commitment was found in the public health interventions. Potential contributions from the social sciences tend to be overwhelmed by the appeal of the biomedical and behavioural sciences. Three concepts and notions notion of community in public health were dominated: First, community- lots and lots of people, or community as the population; second could be described a community as "giant reinforcement schedule" or community as setting, with aspects of that setting being used as levels to support and maintain individual behaviour change. The third, newest, approach sees community as "eco-system with capacity to work towards solutions to its own community identified problems" or to see it as a social system. This notion of community focuses on strengths instead merely on deficits. Two groups of research activities (systematic study of communities and inequality research) supported with evidence from many applied researches were done through development of European Healthy Cities Project and have contributed to this shift in perception of the value of the community. In this course we elaborate inequity research, "System" study of communities and present the case study: "Community applied research in Croatia- "triggered" by Healthy Cities".
Teaching methods	- lectures - seminar presentations and discussion (for the selected topics - each student) - individual/small group seminars paper and presentation preparation
Specific recommendations for teachers	A total of nine teaching hours consist of: Five contacts hours – lectures (two) + seminar presentations and discussion (three) Four individual/small group hours - seminars paper and presentation preparation
Assessment of Students	Seminar paper – selected topics for individual tasks and presentations + Structured essay with selected topics covering most of the course objectives.

COMMUNITY HEALTH - PUBLIC HEALTH RESEARCH METHODS AND PRACTICE

Selma Sogoric, Aleksandar Dzakula

Theoretical background

Community and public health

Throughout the sixties, seventies and eighties, much of the rhetoric in public health paid lip service to the value of a community, empowerment, community-based care, population-based needs assessment and so on, but we could not see much of the evidence of this commitment in the day-to-day service provision of practitioners or in design applied in public health interventions. Potential contributions from the social sciences tend to be overwhelmed by the appeal of the biomedical and behavioural sciences. The most common notion of community in public health was the most simple – lots and lots of people or community as the population. This notion is illustrated in large-scale community interventions propelled by the concern to reach as many people as possible and make best use of scarce program resources. The outcome evaluation of these interventions usually amounts to summing up changes made by individuals in relation to the problem of interest. The greater the number of people who change, the more successful is the intervention (1). The second approach to community borne out of the first could be described as community as "giant reinforcement schedule" or community as setting, with aspects of that setting being used as levers to support and maintain individual behaviour change. In this approach, organizations, groups and key individuals in the community are valued because of their capacity to translate the health messages of the campaign into the local culture. The evaluation of this model rests principally on aggregating changes made by individuals in the population (1-4). The third, newest, approach developed throughout the nineties sees community as "eco-system with capacity to work towards solutions to its own community identified problems" or to see it as a social system. This notion of community focused on strengths instead merely on deficits. The evaluation in these case attempts to capture changes in community processes and structures, as outcomes (1).

Two groups of research activities (systematic study of communities and inequality research) supported with evidence from many applied researches done through development of European Healthy Cities Project contributed to this shift in perception of the value of the community.

Inequality research

Firstly, we acknowledge that people do not live in vacuum. The notion that behaviour is greatly influenced by social context in which people lead their lives has finally get through to public health practitioners. Many sociologists have argued that the lives of individuals are affected not only by their personal characteristics but also by characteristics of the social group they belong (5,6). They say that "lifestyle" and "behaviours" were regarded as matters of free individual choice and dissociated from the social context that shape and constrain them (7-9). With their work, they confronted the prevalent "web of causation" model and blamed it for progressive "individualization" of risks (i.e. attributing risk to characteristics of individuals rather than to environmental or social influences affecting populations). Simultaneously, tremendous shift in value position, from victim blaming, through relative status in social milieu, to density of links and caring in a social structure, has been done through the evolution of causal explanations of inequalities (6).

The traditional explanation of inequalities in health is that they are caused by the behaviour of those from the lower socio-economic classes who drink, smoke and generally engage in too many "risk behaviours" leading to their early demise from heart disease, lung cancer and so on. The solution is therefore, to modify their risky behaviours and so anti-smoking campaigns and other health promotion programs were launched. Unfortunately, historical data shows that such inequalities are independent of the causes of death and they are as prevalent now as they were when the main causes of death were entirely different at the run of the past century.

The next level of explanation is that the inequality is caused by the material deprivation suffered by those in the lower socio-economic groups – poor housing, poor nutrition, inadequate heating, air pollution, inadequate access to care and so on. The solution, therefore, is to provide income or other resource support to the poor in society, enough to raise them above some declared level of deprivation. Although there is undoubtedly some truth to this proposed causal model, Marmot's data from the British civil service study tells us that it is far from the whole story (10). Across five classes of civil servants all of whom are "well-off", there are marked inequalities in health; none suffer what could be called "deprivation". So, aside from the evidence on absolute deprivation, there is growing evidence that the relative distribution of income in a society matters in its own right for population health.

This next level of explanation was given by Wilkinson (11). In his research he found strong negative association between the degree of income inequality in a country and its health as measured by mortality statistics. Here the model proposed is that the feelings of relative deprivation among those in the lower half of the income distribution express themselves through neuro-immunological systems as disease and death. The

larger the differences the more likely and the more severe are the negative health consequences. Low control, insecurity and loss of self-esteem are among the psychosocial risk factors known to mediate between health and socioeconomic circumstances. Exposure to chronic mental and emotional stress (associated with social position) will increase probability of acquiring risky behaviours - stress related smoking, drinking, eating "for comfort", etc. The implied solution in this case would be development of more egalitarian society e.g. the reduction in income inequalities by better distribution of wealth in society.

This level of explanation has been pushed one step further by work of Kennedy and Kawachi (12-14). In the American study by Kennedy income inequality at the state level was strongly correlated with total mortality. Income inequality was measured in that study by the Robin Hood index, which is the proportion of aggregate income that needs to be redistributed from the rich to the poor so as to achieve equality of income. A 1% rise in the Robin Hood index was associated with an excess mortality of 21.7 deaths per 100 000, suggesting that even a modest reduction in inequality could have an important impact on populations health. The maldistribution of income was related not only to total mortality but also to infant mortality, homicides, and deaths from cardiovascular diseases and neoplasm. In an independent study, Kaplan (15) examined the association between income inequality – as measured by the share of aggregated income earned by the bottom 50% of households – and state level variations in total mortality. A strong association was found between their measure of income inequality and age-adjusted total mortality rates in 1990.

Moreover, the degree of income inequality in each state in 1980 was a powerful predictor of levels of total mortality 10 years later. The repeated corroboration of the hypothesis that income inequality is harmful to health has spurred the search for the pathways and mechanisms underlying this relation. One hypothesis was that rising income inequality results in increased level of frustration, which may have deleterious behavioural and health consequences. Societies that permit large disparities in income to develop also tend to be the ones that under invest in human capital (e.g. education), health care, and other factors that promote health. The growing gap between the rich and the poor has led to declining levels of social cohesion and trust, or disinvestments in "social capital". Social capital has been defined as the features of social organization, such as civic participation, norms of reciprocity, and trust in others that facilitate cooperation for mutual benefit. Social capital is thus a community- level variable whose counterpart at the individual level is measured by person's social networks. The core concept of social capital, according to its principal theorists (Putnam) consists of civic engagement and levels of mutual trust among community members. So, by connecting levels of civic trust (perceived levels of fairness and helpfulness) and density of associational membership with degree of income inequity on one side and mortality on other Kawachi isolated social capital, as a mediating mechanism.

The work of Kaplan, Kennedy, and Kawachi is telling us that the growing gap between the rich and the poor affects the social organization of communities and that the resulting damage to the social fabric may have profound implications for public's health. Although the role of economic characteristics in relationship between social capital and health has not been thoroughly elucidated (Veenstra 16) contemporary public health tend to focus less on the individual and more on the social system's influence on health accepting that "the way we organize our society, the extent to which we encourage interaction among the citizenry and the degree to which we trust and associate with each other in caring communities is probably the most important determinant of our health" (6).

"System" study of communities

During the last few decades' significant work has been done by social scientists engaged in the systematic study of communities. From this work, at least two important principles can be identified (17). The first relates to definitions of community, and the second applies a "system" perspective to communities. The first approach to community is based on notion that communities form a whole greater than the assemblage of individuals within them. The community components include locality, an interdependent social group, interpersonal relationship, and a culture that includes values, norms, and attachments to the community as a whole as well to its parts. The second, system view sees communities, simply as a system, which includes individuals, subsystems, and the interrelationship among the subsystems. Anthropologists have identified important subsystems of any community system: political sector, religious sector, recreational sector, and social welfare sector. In addition, community organization studies have identified two additional sectors being important for achieving changes in the community system: voluntary and civic groups, such as health-related agencies, political action groups, and other grass-roots groups, and other groups that may be specific to particular community. From a system perspective, a change in one sector usually implies that adjustment or response will eventually occur in other parts of the system. Change that begins with one sector, however, may take a long time to affect the entire system. In addition, many factors may interrupt of divert the change effort. From a community organization perspective, the target of change is generally the entire system – the community itself. From this perspective, it is not enough to change only a sector or part of the community, although changes in the sectors or subsystems, especially the political or economic spheres, may contribute to overall system change. Sanders (17) delineate the following community components: economic institutions, local government, health, education, social welfare, religion, recreation, social networks, the family and social groupings. Each component could be subdivided and number of community units may be expanded endlessly. Rothman (18)

defines social participation as a core element of social health and for him socially healthy community is the one sufficiently endowed with a matrix of social units that allow participation. In his opinion policy-making unit (decision-making bodies) are critical components of the entire community system, because improving the social or physical health of any population group may require health promotion policies. Those policies could (1) help individuals change their own personal health related behaviours, (2) reduce environmental obstacles to health-promoting behaviour and/or (3) reduce or eliminate factors in the physical or social environment that are detrimental to health. Conscious individual action can be facilitated by "micro-strategies" like for example teaching low income people to prepare nutritious meals using inexpensive food. Often, such individual behaviours are discouraged by environmental conditions such are high prices for more nutritious foods. Introducing food stamps was kind of policy designed, for example, to reduce economic (environmental) barriers to nutritious diet. Finally, environmental hazards may be eliminated by "macro-strategies" e.g. "technological bypass" – by changing potentially detrimental experience of individuals without their direct involvement, like for example legislative action that could lower pricing of more nutritious as compared to less nutritious food. Decision-making units may be favourable, neutral, or unfavourable to establishing relevant policies on the issue. If unfavourable, considerable community effort may be required, a wide scope of community units may need to be mobilized and political pressure tactics may be necessary.

Even if the decision-making body is favourable, substantial community organization may be required as resources are always scarce and competing claims inundate those with decision-making power and responsibility. The source of initiation of the policy change (for example relatively powerless citizens, elite group, or established health professionals or organization) may shape the form of community action.

Rothman (18) described three general forms of community intervention: locality development, social planning, and social action. The first one maximizes local participation (ownership); the second emphasizes rational planning and problem solving and the third uses mobilization and activation of disadvanced groups.

As a part of the effort to influence local public policy community has to get organized. Brown's model of community organization for action (19) comprises four phases – pre-organizational conditions, community organization, policy influence and policy decision. First phase, a pre-organizational condition includes: needs, predisposition to organize and enabling resources (factors) for organizing. Second phase, community organization includes: process of organized action, technical support and expanded (outside the constituency) support and opposition. Third phase, policy influence identifies and described the target of community action: receptivity of policy-making body and noncommunity (external) factors. And fourth phase, policy decision represents the culmination of the whole process, the extent to which policies have been changed according to objectives of the initial action. By offering nine categories of indicators to describe community organization's efforts to influence local public policy Brown's model is helping us to develop better understanding of mechanisms of community action, enables monitoring of process of change and teach us what forms and strategies of community action may be most effective in promoting health.

Another very useful concept for public health practitioners is the one developed by Partick and Wickizer (17). They explained that the social system in a community relevant to health consists of at least three elements: physical structure, social structure and social cohesion. A community's physical structure (urban planning, the design of suburban housing developments, parks and green areas, industry) has both direct influences on health through exposure to risk and indirect influences on health through the creation or neglect of health-inducing environments. Social structure in a community is reflected in such things as its meeting places, mechanisms for income redistribution, sports leagues, clubs, associations and all the elements of a community that allow for the exchange of views and values and engender mutual trust. This, too, has both, direct effects on health, and indirect through facilitating collective problem solving or collective identity. Finally, social cohesion is very much the product of the adequacy of physical and social structure in a community. Along with such things as the cultural or social homogeneity of a community, physical and social structure can either encourage or discourage mutual support and caring, self-esteem and sense of belonging, and enriched social relationship. All of these have been shown, largely by social scientists, to have an influence on the health of a community's members.

Case study

Community applied research in Croatia - "triggered" by Healthy Cities

During more than twenty years of the Healthy Cities project existence in Europe much of the earlier mentioned "theory" has been learned experientially. The Healthy Cities (HC) Project, initiated by the WHO European Office in 1986, is a long-term international development project that seeks to put health on the agenda of decision-makers in cities and to build a strong lobby for public health at the local level. The crucial notion that stimulates HC project development was the recognition of importance of the political will. The Healthy Cities Project challenges cities to take seriously the process of developing health–enhancing public policies that create physical and social environments that support health and strengthen community action for health. Initiating the Healthy Cities Project process requires explicit political commitment and consensus across party political lines, leading to sound project infrastructure, clear strategy, participation mechanisms and broadly-

based ownership (20,21). Healthy Cities is about change, openness to participation, innovation and formal system reorientation. It is changing the ways in which individuals, communities, private and voluntary organizations and local governments think about, understand and make decisions about health.

European cities in general are challenged with complex public health issues like poverty, violence, social exclusion, pollution, substandard housing, the unmet needs of elderly and young people, homeless people and migrants, unhealthy spatial planning, the lack of participatory practices, and unsustainable development (21). Due to the war and post-war transition, Croatian cities are faced with many others, like, for example, mental health, posttraumatic disorders, quality of life of disabled, family health, community regeneration and community capacity building, unemployment, especially among young and mid career workers, stress, alcohol, tobacco and substance misuse, etc (22). The Healthy Cities Project framework provided the testing ground for applying new strategies and methods for addressing these issues in Croatia.

In the early 1990s, migrations caused by war undermined the credibility of the dominant positivist perspective of demographic analyses, statistical studies and quantitative health indicators (23-25). All health indicators obtained at that time were based on estimates of a key factor – population. Quantitative data collected by national health institutions: Croatian Institute of Public Health, Croatian Health Insurance Institute and Ministry of Health, mainly produced mortality and morbidity statistics, which was of some use only for national health policy makers (26-28). In addition to its dubious credibility, national health statistics had other shortcomings: poor accessibility of indicators at the local level and non-inclusion of the opinion of the community (22).

Due to post-war conditions, scarce assets, and the need to determine the state of affairs and launch the action as soon as possible, the method of rapid appraisal (29,30) was chosen for the community health needs assessment and development of the strategic city health documents: the City Health Profile and City Action Plan for Health (31). The most popular and most used method in the Croatian cities is the method of Rapid Appraisal to Assess Community Health Needs (29,30). It was used in 13 cities between 1996 and 2010 (Pula, Metkovic, Rijeka, Karlovac, Varazdin, Zagreb, Split, Dubrovnik, Crikvenica, Slatina, Porec, Labin, Vinkovci). The advantages of this method in comparison with classical approaches to health assessment are as follows: it can be done quickly (in two months from the start), it does not take too much expert time and financial resources (approximately 6.500 EUR per city), it is participatory (representatives of different groups of citizens participate in the process, from needs identification to solution finding; includes representatives of city authorities, institutions and organizations as well as those from non-governmental and non-for-profit sector), sensitive (ability to reflect local particularities), valid (scientifically sound), action-oriented (as a product it gives short-term and long-term plan), and its achievements are sustainable (it establishes and facilitates co-operation among key stake-holders in the project via priority thematic groups).

Academic credibility of described needs assessment method was strengthened by the establishment of strict selection rules of participants and panellists and by the process of triangulation of both information sources (essays, observations and collected objective indicators from the system) and researchers (experts of three different backgrounds: public health, epidemiology and medical information science).

Qualitative analytical approach also was used in development of the model of rapid appraisal of effectiveness of public health interventions (32). A retrospective study of 44 successfully performed interventions in five cities – Liverpool (UK), Sandwell (UK), Vienna (Austria), Pula (Croatia) and Rijeka (Croatia) – identified the indicators of intervention effectiveness that could be used to asses the effect of an intervention in a short period of time (within a time frame of 1-5 years from the beginning of the intervention) by measuring several aspects of success. These are as follows:

1. Effect on political environment (macro-environment) – assessment of the achieved degree of change in political environment;
2. Effect on a project user – an individual, a group, a community, within the meaning of empowering users and influencing health;
3. Effect on a project manager – an organization or institution, i.e., an association or group (microenvironment); and
4. Monitoring the effectiveness of the implementation process of an intervention.

The instrument happened to be more applicable for measuring the success of individual (population- or topic-targeted) interventions. In the evaluation of effectiveness of comprehensive years-long interventions, such as Healthy City or Healthy County, it is applied together with other evaluation instruments.

Concluding remarks

The job of public health professionals, including those in academic setting, is not only to investigate and understand the world; it is also to change it. This is why we, in Croatia put the emphasis on the development of applied (action) research by which the academic knowledge may be used for intensifying activities and development of local communities.

The introduction of participatory methods and consensus building techniques in the process of public health policy formulation in Croatia has brought much better understanding and improved collaboration among "policy stakeholders" - politicians, administration, public health professionals and community. Public health professionals are more responsive and committed to work with communities to support them to (re)generate local social capital. Croatian Healthy Cities and Counties greatest achievement, at the moment, is that community participation is assured in all stages of planning and management of the resources for health at the local level.

References

1. Hawe P. Capturing the meaning of "community" in community intervention evaluation: some contributions from community psychology. Health Prom Int 1994;9:199-210.
2. Koepsell TD, Diehr PH, Cheadler A, Kristal A. Invited commentary: symposium on community intervention trials. Am J Epidemiol 1995;142:594-9.
3. Feinleib M. New directions for community interventions studies. Am J Public Health, 1996;86:1696-8.
4. Murray DM. Design and analysis of community trials: lessons from the Minnesota Heart Health Program. Am J Epidemiol 1995;142:569-75.
5. Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. Am J Public Health 1998;88:216-22.
6. Lomas J. Social capital and health: implications for public health and epidemiology. Soc Sci Med 1998;47:1181-8.
7. Hawe P, Shiell A. Social capital and health promotion: a review. Soc Sci Med 2000; 51:871-85.
8. McKinlay JB. The promotion of health through planned sociopolitical change: challenges for research and policy. Soc Sci Med 1993;36:109-17.
9. Shy CM. The failure of academic epidemiology: witness for the prosecution. Am J Epidemiol 1997;145:479-84.
10. Marmot MG, Rose G, Shipley M, Hamilton PJ. Employment grade and coronary heart disease in British civil servants. J Epidemiol Community Health 1978;32:244-9.
11. Wilkinson RG. Socioeconomic determinants of health. Health inequalities: relative or absolute material standards? BMJ 1997;314:591-5.
12. Kennedy BP, Kawachi I, Prothrow-Stith D. Income distribution and mortality: cross-sectional ecological study of the Robin Hood Index in United States. BMJ 1996;312:1004-7.
13. Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality and mortality. Am J Public Health 1997;87:1491-8.
14. Kawachi I, Kennedy BP. Health and social cohesion: why care about income inequality? BMJ 1997;314:1037-40.
15. Kaplan GA, Pamuk ER, Lynch JW, Cohen RD, Balfour JL. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. BMJ 1996;312:999-1003.
16. Veenstra G. Location, location, location: contextual and compositional health effects of social capital in British Columbia, Canada, Soc Sci Med 2005;60:2059-71.
17. Thompson B, Kinne S. Social change theory. Applications to community Health. In:Bracht N, editor. Health promotion at the community level. London: Sage Publications;1990. p.45-65.
18. Rothman J, Brown ER. Indicators of societal action to promote social health. In: Kar SB, editor. Health promotion indicators and actions. New York: Springer; 1996. p. 202-20.
19. Brown ER. Community organization influences on local public health care policy: a general research model and comparative case study. Health Educ Q 1983;10:205-33.
20. WHO Regional Office for Europe. Centre for Urban health. Strategic Plan. Urban Health/ Healthy Cities Programme (1998 – 2002). Phase III of the WHO healthy Cities project. Copenhagen: 7 June 1998.
21. WHO Regional Office for Europe. Center for Urban health. WHO Healthy Cities Project Phase III: 1998 – 2002 Terms of reference for the WHO Healthy Cities network. Copenhagen: 14 January 1999.
22. Sogoric S, Vukusic Rukavina T, Brborovic O, Oreskovic S, Zganec N, Vlahusic A. Counties Selecting Public Health Priorities– A "Bottom - Up" Approach (Croatian Experience). Coll. Antropol. 2005;29(1):111-119.

23. Bozicevic I, Oreskovic S, Stevanovic R, Rodin U, Nolte E, McKee M. What is happening to the health of the Croatian population? Croat Med J 2001;42:601-5.
24. Levett J. On the humanitarian disaster in the Balkans from a public health perspective. The Internet Journal of Public Health Education 1999;1:1-9.
25. Lang S, Kovacic L, Sogoric S, Brborovic O. Challenge of goodness III: public health facing war. Croat Med J 2002;43:156-65.
26. Turek S, Rudan I, Smolej-Narancic N, Szirovicza L, Cubrilo-Turek M, Zerjavic-Hrabak V et al. A large cross-sectional study of health attitudes, knowledge, behaviour and risks in the post-war Croatian population (the First Croatian Health Project). Coll Antropol 2001;25:77-96.
27. Turek S. Reform of health insurance in Croatia. Croat Med J. 1999;40:143-51.
28. Hebrang A. The direction of Croatian health care in the post-war period [in Croatian] Lijec Vjesn 1991;113(9-10):301-8.
29. Sogoric S. Application of the Modified Method of "Rapid Appraisal to Assess Community Health Needs" for making Rapid City Health Profiles and City Action Plans for Health. Croat Med J 1998;39:267-75.
30. Sogoric S, Vukusic Rukavina T, Mujkic A, Dzakula A. Formulating a health strategy at the local level using the "Urban Health Profile" and "Urban Health Planning" [in Croatian] Lijec Vjesn 2003;125(11-12):334-8.
31. Garcia P, McCarthy M. Measuring health. A step in development of city health profiles. Copenhagen: WHO; 1996.
32. Sogoric S, Middleton J, Lang S, Ivankovic D, Kern J. A naturalistic inquiry on the impact of interventions aiming to improve health and the quality of life in the community. Soc Sci Med. 2005 Jan; 60(1):153-64.

Recommended readings

1. Community participation in local health and sustainable development. Approaches and techniques. http://www.euro.who.int/document/e78652.pdf.
2. Social Determinants of Health: the Solid Facts. Second edition. http://www.euro.who.int/__data/assets/pdf_file/0005/98438/e81384.pdf.
3. Health 2020: a European policy framework supporting action across government and society for health and well-being http://www.euro.who.int/__data/assets/pdf_file/0009/169803/RC62wd09-Eng.pdf.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Economic evaluation in health care: Practical approach**
Module: 2.37	ECTS (suggested): 0.2
Author(s), degrees, institution(s)	Pia Vracko, MD, PhD Candidate, Research Assistant, Teaching Assistant National Institute of Public Health of the Republic of Slovenia, Slovenia Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	Pia Vracko National Institute of Public Health of the Republic of Slovenia, Trubarjeva ulica 2, 1000 Ljubljana, Slovenia E-mail: pia.vracko@ivz-rs.si
Keywords	Cost-benefit, cost-effectiveness, cost-minimization, cost-utility, economic evaluation, health economics.
Learning objectives	After completing this module students and public health professionals should be able to: • understand economic evaluation terminology; • describe methodology of cost analysis and find data for it; • describe methodology and understand different uses of cost-minimization, cost-effectiveness, cost-utility, and cost-benefit analyses; • understand the importance of certain issues that are important in economic evaluation in healthcare; • independently interpret the economic evaluation studies.
Abstract	Economic evaluation of healthcare interventions is a useful tool in resource allocation planning. It can enumerate the potential costs and value the anticipated consequences of alternative programs, and analyze risks and uncertainties for their successful outcomes. In this module, four different types of economic evaluation are presented: cost-minimization analysis, cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis.
Teaching methods	An introductory lecture gives the students insight into the characteristics of economic evaluation techniques. The theoretical knowledge is illustrated by case studies. In continuation, students should distinguish among four types of economic evaluation, their value, and limitations in decision-making process. In group work, students first perform all four different types of economic evaluation in a simple example, and than present self-selected cases of four types of economic evaluation.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple-choice test and calculation of various measures of economic evaluation in practice.

ECONOMIC EVALUATION IN HEALTH CARE: PRACTICAL APPROACH

Pia Vracko, Lijana Zaletel-Kragelj

Theoretical background

What is an economic evaluation?

Economic evaluation may be defined as the comparative analysis of alternative courses of action in terms of both their costs and consequences (1). Based upon welfare economics, a subfield within economics that has explored the question of how to maximize consumer welfare within a limited budget, it tries to assess the social desirability of a program relative to some other alternative. Economic evaluation can enumerate the potential costs and value the anticipated consequences of a proposed intervention, program, policy or regulatory initiative, and reflect trade-offs in alternatives. The additional expense of the new intervention means that resources have to be redirected from elsewhere and an economic evaluation assesses whether or not the additional benefits generated by the new intervention are greater than the loss in benefits from the reduction in other programmes - that is, is the reallocation efficient? Additionally, an economic evaluation can analyze risk and uncertainty for a successful outcome of a project (2).

The ultimate goal of economic evaluation in healthcare is to maximize net health benefits for all persons in a target population given a range of health care interventions and known resource constraints (3). With the healthcare sector accounting for a sizeable proportion of national expenditures, the issue of efficiency has become a central objective of policymakers within most health systems.

What does economic evaluation involve?

The basics of economic evaluation involve identifying, measuring, valuing, and comparing the costs and consequences of alternatives being considered (1).

The measurement of costs is always in monetary units. The outputs, on the other hand, can come in different forms, and correspondingly there are different types of economic evaluation that are described in next section.

In addition to costs and consequences, a reliable economic evaluation should involve:

1. Clear definition of the viewpoint from which the analysis is performed; an item may be a cost (or consequence) from one point of view, but not from another. For example, patients' travel costs are a cost from the patient's point of view and from the society's point of view, but not a cost from the Ministry of Health's point of view. Possible points of view include those of society, the Ministry of Health, other government ministries, the government in general, the patient, the employer, and the agency providing the program (1). Optimally, the analyst should adopt the societal point of view, which is the broadest one and is always relevant;
2. Time structure of intervention; usually, costs and consequences do not occur at the same time, therefore costs (and benefits in cost-benefit analysis) should be discounted to net present value;
3. Analysis of risk and uncertainty (of successful outcome); uncertainty in economic evaluation can arise because of methodological disagreement among analysis, the data requirements of the study, the need to extrapolate results over time, or the desire to generalize the results of the study to another setting (4);
4. Distributional considerations (distribution of wealth in society, equity), which markets ignore and the healthcare sector considers an essential ingredient.

For in-depth information on theoretical background of economic evaluation, please refer to the module by Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems (5).

Cost analysis

Cost analysis is used in all four types of economic evaluation. In healthcare field the costs are disaggregated into three categories: direct, indirect and intangible. Direct means directly related to healthcare industry (the stuff, the hospitals and the patients), indirect refers to inputs and outputs outside healthcare industry, and intangible refers to the pain and suffering that are caused or alleviated by a healthcare intervention. For example, physician and nursing expenses are direct costs, and loss of income due to illness is referred to as indirect cost.

According to Drummond et al. (1), steps in cost analysis are the following:

1. Identification of resource categories, for example:
    a. health care resource use
        - hospital resources (investigations, therapy, bed days, out-patient attendances, overheads)
        - community care resources (general practitioner visits, nurse visits, ambulance or hospital car)

b. patient and family resource use (patient's time, time of relatives, out-of-pocket expenses for transport)
c. resource use in other sectors (social workers visits, home help visits).

2. Measurement of resource use
   The most accurate data collection is by clinical trial, as case report forms are completed for patients enrolled in the trial. Normally these record data on clinical events, but they can be modified to include resource use, such as number and type of investigations, date of hospital admission and discharge. In the absence of a trial the two major sources of data on resource use are routine statistics kept at the hospital or by other agencies, and patient's case charts.
3. Valuation of resource items
   This is dependent on the availability of local financial data. In some settings, there may be data on hospital billings or charges, usually kept by agencies. In other settings, detailed costing studies would be necessary.

Types of economic evaluation

There are four types of economic evaluation (1,6). One type only uses costs; it is called a cost-minimization analysis. The other three types differ according to what kind of consequence they incorporate along with costs.

The most obvious consequence is what the healthcare industry immediately works with, such as diagnostic test outcome or an operation successfully completed. These outputs are called "effects", and expressed in natural units (such as percentage detection).

A broader measure of effects relies on "utilities" (i.e., estimates of the satisfaction of the effects), and the output unit is called a "quality-adjusted life-year", (QALY). The QALY is a measure of the value of health outcomes. Since health is a function of length of life and quality of life, the QALY was developed as an attempt to combine the value of these attributes into a single index number. The QALY calculation is simple: the change in utility value induced by the treatment is multiplied by the duration of the treatment effect to provide the number of QALYs gained. QALYs can then be incorporated with medical costs to arrive at a final common denominator of cost/QALY. This parameter can be used to compare the cost-effectiveness of any treatment (7).

Lastly, the output can be expressed in the same monetary unit as the costs, in which case the consequences are called "benefits".

"Effects" are used in a cost-effectiveness analysis, "QALYs" are used in cost-utility analysis, and "benefits" are used in cost-benefit analysis (8).

Cost-minimization analysis

Cost-minimization analysis (CMA) is used to compare net costs of healthcare interventions with exactly the same effects (quality and quantity of the effect are the same). The costs of given healthcare interventions are analyzed and compared. The aim is to choose the intervention that supplies an effect at lowest cost. Under a CMA, treatment 1 would be more cost-efficient (that is, less costly) than treatment 2 if (Equation 1):

$$C_1 < C_2 \qquad \textbf{Equation 1.}$$

C_1 = cost of treatment 1
C_2 = cost of treatment 2

If we want to calculate the difference between costs of treatment 1 and costs of treatment 2 we use the simple procedure (Equation 2):

$$\Delta C = C_1 - C_2 \qquad \textbf{Equation 2.}$$

ΔC = *cost difference*
C_1 = cost of treatment 1
C_2 = cost of treatment 2

To summarize, a CMA identifies, measures and compares only input costs of alternative interventions what is, due to its simplicity, an advantage. But, it assumes outcomes to be equivalent. It requires clinical evidence that differences in health effects between alternatives are minimal or not important. The major problem is that the assumptions are in most cases difficult to justify prior to any experimental study. This method, therefore, has only limited application what is its disadvantage. A common example of a CMA is the comparison of drugs that are of the same chemical entity, the same dose, and have the same pharmaceutical properties as each other (e.g. innovative/original drug versus generic drug). In this case only the cost of the drug itself needs to be compared because outcomes should be the same.

Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) is used to compare net costs of healthcare interventions per unit of the effect. A single effect common to the given alternatives, but achieved to different degrees, is considered and expressed in natural clinical units, such as years of life saved, premature deaths avoided or improvements in functional status (units of blood pressure or cholesterol).

The aim is to choose the intervention that supplies a unit of effect at lowest cost. or one can try to achieve the most effect per monetary unit (i.e. euro, dollar etc.) of cost. Under a CEA, treatment 1 would be more cost-effective that treatment 2 if (Equation 3):

$$\frac{C_1}{E_1} < \frac{C_2}{E_2} \qquad \textbf{Equation 3.}$$

C_1 = cost of treatment 1
C_2 = cost of treatment 2
E_1 = units of health effects achieved by treatment 1
E_2 = units of health effects achieved by treatment 2

Comparing the costs and health outcomes of 2 treatments' results in 1 of 9 pairings (Figure 1). Since costs and benefits are measured in non-comparable units, their ratio (Equation 2) provides a yardstick with which to assess relative (productive) efficiency (9). It does not, however, enable us to evaluate the relative efficiency of interventions which provide more benefit at greater cost or less benefit at lower cost (10).

For decision a cost-effectiveness grid can be used (6). In the case the health outcomes are equivalent (Figure 1, centre column), the choice should be based on cost; when costs are equivalent (Figure 1, centre row), the choice should be based on outcome. When one strategy has better outcomes and lower costs (Figure 1, upper right and lower left boxes), the choice is clear. The decision is difficult only when the strategy that is more expensive also produces better outcomes (Figure 1, upper left and lower right boxes).

Figure 1. Cost-Effectiveness Comparison of Treatments 1 and 2 and Possible Decisions (6). LEGEND: C=costs, E=effect, 1=treatment 1, 2=treatment 2

	HEALTH OUTCOME		
COST	$E_1 > E_2$	$E_1 = E_2$	$E_1 < E_2$
$C_1 > C_2$	Calculate incremental cost-effectiveness ratio	2 less expensive Choose 2	2 dominates 1 Choose 2
$C_1 = C_2$	1 better effect Choose 1	Makes no difference	2 better effect Choose 2
$C_1 < C_2$	1 dominates 2 Choose 1	1 less expensive Choose 1	Calculate incremental cost-effectiveness ratio

If an intervention is both more expensive and more effective than an alternative, then the criterion for efficiency becomes the incremental cost effectiveness ratio, iCER. It is defined as the ratio of the change in costs of two programmes to the change in effects of the programmes (Equation 4):

$$\text{iCER} = \frac{C_1 - C_2}{E_1 - E_2} \qquad \textbf{Equation 4.}$$

iCER = incremental cost-effectiveness ratio
C_1 = cost of treatment 1
C_2 = cost of treatment 2
E_1 = units of health effects achieved by treatment 1
E_2 = units of health effects achieved by treatment 2

iCER represents the additional cost of an intervention divided by the additional health outcome it achieves; in fact, iCER provides information on how much each additional therapeutic success will cost. It is most useful when a choice must be made among several therapeutic interventions. The lower the ratio, the greater is the health improvement for a given resource expenditure. A ratio increases most rapidly as its denominator approaches 0 (this is, when the 2 interventions provide nearly equal health effects).

A major limitation of cost-effectiveness analysis is its inability to compare interventions with differing natural effects (1). For example, interventions aimed at increasing life years gained cannot be directly compared with those which improve physical functioning. Cost-effectiveness analysis therefore cannot directly address resource allocation among alternative, unrelated health interventions (10).

Cost-utility analysis

Cost-utility analysis (CUA) is used to compare net costs of healthcare interventions with different effects, when prices do not want to be used explicitly. Therefore, all effects need to be converted to a common unit, a quality adjusted life year (a QALY). The aim is to choose the intervention that supplies a unit of effect, in terms of QALYs, at lowest cost. In CUA, the following criterion is used to replace Equation 2 (Equation 5):

$$\frac{C_1}{QALY_1} < \frac{C_2}{QALY_2} \qquad \textbf{Equation 5.}$$

C_1 = cost of treatment 1
C_2 = cost of treatment 2
$QALY_1$ = QALYs achieved by treatment 1
$QALY_2$ = QALYs achieved by treatment 2

In cost-utility analysis, relative efficiency is assessed using an incremental cost-utility ratio (iCUR). It is defined as the ratio of the change in costs of two programmes to the change in utility of the programmes (Equation 6).

$$iCUR = \frac{C_1 - C_2}{QALY_1 - QALY_2} \qquad \textbf{Equation 6.}$$

iCUR = incremental cost-utility ratio
C_1 = cost of treatment 1
C_2 = cost of treatment 2
$QALY_1$ = QALYs achieved by treatment 1
$QALY_2$ = QALYs achieved by treatment 2

In CUA, the use of a single measure of health benefits enables the comparison of the efficiency of resource allocation among alternative, unrelated health interventions. The optimal decision rule involves ranking the incremental cost-utility ratios of different interventions and selecting those with the lowest ratios (best value) until the budget is depleted (11,12). The lower the incremental ratio, the higher the priority in terms of maximizing health benefits derived from a given level of expenditure (12).

The value of CUA is limited by understanding and meaningful interpretation of QALYs. Individuals are not used to buying a QALY and are therefore unfamiliar with the process of trying to obtain a QALY at a lowest cost. Deriving meaningful estimates of QALYs is therefore at the heart of the CUA evaluation.

Cost-benefit analysis

Cost-benefit analysis (CBA) places a monetary value on the consequences of each alternative health interventions. Forming 'benefits' means that the consequences are measured using the same unit as 'costs'. One can tell whether the benefits are greater than costs and thus know whether the expenditure is worthwhile. The basic cost-benefit criterion is (Equation 7):

$$B_1 > C_1 \qquad \textbf{Equation 7.}$$

B_1 = benefits of treatment 1
C_1 = cost of treatment 1

Benefits can also be expressed as a product of effect (E) and price (P). In this case, the Equation 6 can be rewritten as (Equation 8):

$$P_1 \times E_1 > C_1 \qquad \textbf{Equation 8.}$$

C_1 = cost of treatment 1
E_1 = units of health effects achieved by treatment 1
P_1 = attributable price of effects achieved by treatment 1

$$\frac{C_1}{P_1 \times E_1} < \frac{C_2}{P_2 \times E_2} \qquad \textbf{Equation 9.}$$

C_1 = cost of treatment 1
C_2 = cost of treatment 2
E_1 = units of health effects achieved by treatment 1
E_2 = units of health effects achieved by treatment 2
P1 = attributable price of effects achieved by treatment 1
P2 = attributable price of effects achieved by treatment 2

When there is a financial budget constraint, using funds for one purpose precludes their use for another. In this case, the CBA criterion has an alternative formulation (Equation 9). This criterion enables to choose the intervention that supplies a unit of benefit at lowest cost (7).

To calculate the difference in cost-benefit between the two (or more) programmes being compared in the evaluation, incremental cost-benefit ratio (iCBR) may be used. It is defined as the ratio of the change in costs of two programmes to the change in benefits of the programmes (Equation 10):

$$\text{iCBR} = \frac{C_1 - C_2}{B_1 - B_2} \qquad \textbf{Equation 10.}$$

iCBR = incremental cost-benefit ratio
C_1 = cost of treatment 1
C_2 = cost of treatment 2
B_1 = benefits of treatment 1
B_2 = benefits of treatment 2

The result of iCBR shows additional euros earned with each additional euro invested in new treatment.

Another useful criterion in CBA is the so called net benefits (NB) criterion (Equation 11):

$$\text{NB} = (B_1 - B_2) - (C_1 - C_2) \qquad \textbf{Equation 11.}$$

NB = net benefits
C_1 = cost of treatment 1
C_2 = cost of treatment 2
B_1 = benefits of treatment 1
B_2 = benefits of treatment 2

The result of this criterion represents the excess of monetary benefits over costs.

Welfare economics shows that under certain conditions any net excess of monetary benefits over costs represents the gain in welfare by society (13). Cost-benefit analysis therefore makes it possible to determine, firstly, whether an individual intervention offers an overall net welfare gain and, secondly, how the welfare gain from that intervention compares with those from alternative interventions. Increased use of interventions with the greatest net gain will increase efficiency. By valuing all costs and benefits in the same units, cost-benefit analysis compares diverse interventions using the net benefit criterion. Cost-benefit analysis thus simultaneously addresses issues of productive efficiency and issues of resource allocation among alternative, unrelated health interventions.

There has been a general unease with the pricing of the effects in CBA. Commonly, the monetary value is attributed to outcomes through 'willingness to pay' approach (i.e. how much patients would be willing to pay for a given health benefit, such as avoidance of pain or disability, or, in public health, how much a health agency would be willing to pay for prevention/treatment of a disease in a population) (14).

Types of economic evaluation in practice

When relevant effects of the given health care interventions are observed to be similar, CMA is used to compare net costs.

When a single effect is common to the alternative interventions, but achieved to different degrees, and expressed in natural clinical units, CEA is used to compare alternative interventions in terms of net cost per unit of health effect obtained.

Many times, economic analysts may want to assess interventions with outcomes that are either common or not common to the alternative interventions. In these situations, a specific measure of value must be applied to the relevant outcomes to allow for relevant comparisons with a common denominator. In these cases, the CUA, with QALYs as a common unit, and CBA with common monetary unit, may be applied.

Case study

Case study 1: Costing alternative radiotherapy treatments

Introduction

Cost analysis is the first and inevitable step in every economic evaluation. A reliable cost analysis may be complex and sometimes challenging due to many different data sources. Many times, cost data are hard to get. Therefore, it is meaningful to represent a comprehensive cost analysis (with suggested data sources) in this case study. Drummond et al. (1) based it on an actual study, undertaken in the UK (15 Economic Analyses). For more detailed directions on costing, please refer to the recommended readings.

Description of a case

A clinical trial is being carried out comparing two forms of radiotherapy for patients with head and neck cancer and carcinoma of the bronchus:

- patients receiving conventional therapy are treated once per day, 5 days per week, for about 6 weeks; they would normally travel on a daily basis to a hospital-based radiotherapy centre to receive care;

- patients receiving continuous hyperfractionated accelerated radiotherapy (CHART) are treated three times on each of 12 consecutive days, including the week-end; because of the intensity and frequency of treatment, patients would normally stay in hospital during therapy, either in a regular hospital ward or in a hostel owned by the hospital.

The different treatment regimens obviously give rise to different costs. However, in addition, there may be differences in the period following treatment for the following reasons:

1. the higher intensity of the CHART regimen might give rise to more side-effects, and hence a greater need for community care after hospital discharge;
2. the CHART regimen might give better tumour control, thereby slowing down the progression of the disease;
3. CHART might reduce the extent of late radiation changes, and a lower incidence of necrosis may also reduce the need for salvage surgery.

The clinical trial will provide an opportunity to gather data on the use of resources by patients in the two treatment groups. You are asked to

1. identify which categories of resource you feel it would be important to assess;
2. indicate how you might measure the use of these resources in physical units;
3. say how you might value the resource consumption in money terms.

Identification of resource categories

Resource use can be considered under the following broad headings: hospital resources, community care resources, patient and family resource use, and resource use in other sectors (Table 1).

Table 1. Identification of resource categories for economic evaluation

	Resource items
Hospital resources	• radiotherapy, • bed days, • out-patient attendances, • overheads
Community care resources	• general practitioner visits, • nurse visits, • ambulance or hospital car
Patient and family resource use	• patient's time, • time of relatives, • out-of-pocket expenses for transport (e.g. car, train, or taxi)
Resource use in other sectors	• social workers visits, • home help visits

Measurement of resource use

The fact that a clinical trial is taking place greatly increases the opportunity for accurate data collection as case report forms are completed for patients enrolled in the trial. Normally these record data on clinical events, but they can be modified to include resource use, such as number and type of investigations, date of hospital admission and discharge. Also, the fact that patients are enrolled in a trial provides the opportunity to interview them about the resource use in community care, time taken to travel to hospital, and personal expenditure. They can also be given diary cards to record expenditure or time spent by relatives in home nursing.

In the absence of a trial the two major sources of data on resource use are routine statistics kept at the hospital or by other agencies, and patients' case notes (charts). The quality of these records varies by agency and data are usually more comprehensive at the mail place (clinic) where the patient is being treated. In addition, there are no routine records for patient and family resource use.

Turning to the specific resource items identified above, we might expect to record quantities used as follows (Table 2).

Valuation of resource items

It is extremely difficult to give general advice on this because it is so dependent on the availability of local financial data. In some settings, like the USA, there may be data on hospital billings and charges. In other settings, detailed costing studies would be necessary. When using charge data, it is important to:

1. investigate the relationship between charges and costs;
2. record physical quantities as well as charges, so as to facilitate generalization of study results to other settings.

Turning to the specific resource items measured above, we might expect to value them as follows (Table 3).

Finally, a few rare events, such as hospital admission for particular types of surgery, may be handled separately. Depending on how quantitatively important they seem, case-mix group costs or disease-specific per diems may suffice. Alternatively, micro-costing may be undertaken.

Table 2. Measurement of resource use in economic evaluation

Resource items	Description
Hospital care	
Radiotherapy	The number of treatment sessions could be recorded, possibly differentiating by length of session and time of day (e.g. normal working hours, after hours, weekends)
Bed days	The number of bed days could be recorded, differentiating by type of hospital ward
Out-patient attendances	The number of attendances could be recorded
Overheads	These would probably be related to the number of bed days or other suitable resource item
Community care	
General practitioner visits	The number could be ascertained, either by asking patients, or by consulting the general practitioners. It may make sense to differentiate between home visits and visits to the practitioner's office
Nurse visits	The number could be recorded as for general practitioner visits above. The purpose of the nurse visit and type of nurse (e.g. general nurse, specialist cancer nurse) would be recorded
Ambulance or hospital car	The number and length of trips could be recorded. Length of trip could be ascertained from the patient's place of residence
Patient and family resources	
Patient's time	The time taken in seeking and receiving care could be estimated by asking the patient. Time off work could be estimated separately
Relatives' time	Relatives could spend time in home nursing and in accompanying patients to hospital. It could be estimated as for patients' time above
Out-of-pocket expenses	Some may be estimated directly in money terms, others may be estimated by asking patients
Resources in other sectors	
Social worker and home help visits	These would be estimated in a similar way to nurse visits above.

Table 3. Valuation of resource items in economic analysis

Resource items	Valuation
Hospital care	
Radiotherapy treatment sessions	In some settings there may be charge data, or average cost figures, for radiotherapy sessions. However, even if these exist, which is unlikely in many locations, they may not differentiate by type of session (for example, normal hours, out-of-hours, or weekend). This distinction is critical to understanding the relative cost of conventional radiotherapy and CHART. Therefore, it is likely that micro-costing would be required. In micro-costing the approach would be to derive the cost of a treatment session from its component parts, namely consultant (medical) time, radiographer time, medical physics time, consumables, equipment, buildings, and departmental overheads. Some survey work may be required, plus data from the hospital finance department on staff salaries, overtime allowances, and equipment prices. Costing of equipment and buildings will require assumptions to be made about useful life and resale value. It would be necessary to apportion these costs to individual treatment sessions. Judgements would also need to be made about which components of hospital overheads (for example, cleaning, building maintenance, or administration) are most appropriately allocated to departments and the allocation basis (for example, square metres, cubic metres, number of stuff, and so on). Some elements of overhead may be better allocated on the basis of in-patient days or number of patients.

Bed days	It may be possible to use the average daily costs (or per diems) for different types of wards, including hostel wards. However, these may be considered too imprecise, in which case micro-costing might be undertaken. This would derive a daily cost for a particular category of ward by considering nurse stuffing levels, medical (consultant) input, and overheads. Because hostel wards may not feature in the standard hospital accounts, micro-costing may be required for these, for example, they may be slightly off site or rely partly on staffing by volunteers. An opportunity cost for volunteer time may have to be inputted. In costing hospital beds it may be decided to make an allowance for the fact that there is usually less than 100% occupancy.
Out-patient attendances	There may be an average cost or charge available for an out-patient visit, although this may not differentiate between oncology and other clinical specialties. Depending on the quantitative importance of this item, micro-costing may be undertaken.
Overheads	As mentioned above, these could be allocated to the radiotherapy treatments, to out-patient attendances, or to hospital bed days, depending on the overhead item.
Community care	
General practitioner visits	There may be data available on physician fees for various types of visits (e.g. general assessment, home visit, etc.). Alternatively, there may be nationally available data on the average costs of various general practitioner services. Failing this, micro-costing may be required. This would calculate the cost of practitioners' time (per minute or per hour) and add the cost of travel for home visits. Drug costs would also need to be considered
Nurse visits	The agencies providing the nurses may have data on the average cost of a visit. This may even distinguish between various types of visit. Failing this, micro-costing would have to be employed, taking into account nursing salaries, length of visits, travel time, and nurses time spent in general administration. There may also be some consumables to be accounted for in the cost of nurse visits.
Ambulance and hospital car	Estimates may be available for the average cost per mile travelled. This could be combined with data on the distances involved to generate total costs.
Patient and family resources	
Patients' time	If the time was taken from work-time, the gross salary (including employment benefits) could be used. Different assumptions could be made about the opportunity cost of leisure time
Relatives' time	In general the valuation of this raises the same issues as the valuation of patients' time. The valuation of time spent in informal nursing care is complicated because the relative may also be able to carry out other tasks at the same time.
Resource items	**Valuation**
Out-of-pocket expenses	In general, the financial expenditures made (for example, bus fares) would suffice. However, for some items, such as use of one's private car, the expenditures (say) on fuel would underestimate the true cost. Here, monitoring organizations can often provide data on the cost (per mile or kilometre) of running a car.

Case study 2: Calculation of econometric measures in practice

Data set

The four main methods of an economic evaluation in health care have been introduced in this module. The study questions that follow require the student to select and assemble the appropriate figures from a set of categories of costs and outcomes to conduct each of the four kinds of evaluation.

The evaluation being considered in Table 4 is based upon fictive data on provision of intensive care that involves increased current capital expenditures in order to increase a patient's future survival chances. The costs and consequences for patients are listed in Table 4.

Table 4. Evaluation of intensive care treatment

Cost or consequence	Before intensive care	With intensive care	Incremental effect*
Cost per additional survivor (to hospital discharge)	3,400€ C_1 (hospital)	12,200€ C_2 (hospital)	8,800€
Cost per additional survivor (to death)	72,500€ C_1 (society)	80,100€ C_2(society)	7,600€
Survival rate (to hospital discharge)	65.4% E_{1a}	80.2% E_{2a}	14.8%
Survival time (per live birth):			
a. Life-years	18.8 E_{1b}	27.7 E_{2b}	8.9
b. QALYs	17.4 $QALY_1$	36.0 $QALY_2$	8.6
Earnings	92,200€ B_1	124,200€ B_2	32,000€

*Incremental effect represents the difference, in cost or effect, between the two programmes being compared in economic evaluation.

Cost-minimization analysis

A CMA comparison of intensive care (before versus with intensive care) from: (a) the hospital's perspective and (b) society's perspective is presented in Equation 11 and Equation 12. For calculation Equation 2 should be used.

a.) Hospital perspective (Equation 12):

$$\Delta C_{(\text{hospital})} = 12{,}200€ - 3{,}400€ = 8{,}800€ \qquad \textbf{Equation 12.}$$

$\Delta C_{(\text{hospital})}$= *cost difference from hospital's perspective*

From the hospital perspective, cost per additional survivor before intensive care is 3.400€ and with intensive care is 12,200€. The intensive care increases the cost per additional survivor for 8,800€.

b.) Societal perspective (Equation 13):

$$\Delta C_{(\text{society})} = 80{,}100€ - 72{,}500€ = 7{,}600€ \qquad \textbf{Equation 13.}$$

$\Delta C_{(\text{society})}$= *cost difference from society's perspective*

From the societal perspective, cost per additional survivor before intensive care is 72,500€ and with intensive care is 80,100€. The intensive care increases the cost per additional survivor for 7,600€.

Cost-effectiveness analysis

A CEA comparison of intensive care from society's perspective reveals that intensive care provides a better health outcome, but it is also more expensive. Consequently, the incremental cost-effectiveness ratio should be calculated (Equation 14 and Equation 15 for survival rate and survival time, respectfully). For calculation Equation 4 should be used.

a.) Survival rate (Equation 13):

$$\text{iCER}_{(\text{society})} = \frac{72{,}500€ - 80{,}100€}{65.4 - 80.2} = \frac{-7{,}600€}{-14.8} = 513.5€ \qquad \textbf{Equation 14.}$$

$\text{iCER}_{(\text{society})}$ = incremental cost-*effectiveness ratio from society's perspective*

From the societal point of view, an increase in survival rate per 1% would cost additional 513.5€

b.) Survival time (Life-years) (Equation 14):

$$\text{iCER}_{(\text{society})} = \frac{72{,}500€ - 80{,}100€}{18.8 - 27.7} = \frac{-7{,}600€}{-8.9} = 853.9€ \qquad \textbf{Equation 15.}$$

$\text{iCER}_{(\text{society})}$ = incremental cost-*effectiveness ratio from society's perspective.*

From the societal point of view, a prolongation of survival time for 1 life-year would cost additional 853.9€.

Cost-utility analysis

A CUA comparison of intensive care from society's perspective is presented in Equation 16. To form CUA comparison, only QALYs can be used as a consequence. For calculation Equation 6 should be used:

$$\text{iCUR}_{(\text{society})} = \frac{72{,}500€ - 80{,}100€}{17.4 - 36.0} = \frac{-7{,}600€}{-18.6} = 7{,}581.4€ \qquad \textbf{Equation 16.}$$

$\text{iCUR}_{(\text{society})}$ = incremental cost-*utility ratio from society's perspective*

From the societal point of view, one QALY per survivor gained would cost additional 7,581.4€.

Cost-benefit analysis

For CBA comparison of intensive care from society's perspective, the iCBR and Net benefit should be calculated. iCBR is presented in Equation 17 and net benefit is presented in Equation 18. To form CBA comparison, only earnings/benefits in monetary terms can be used as a consequence. For calculation Equation 10 and Equation 11 should be used.

a.) iCBR (Equation 17):

$$\text{iCBR}_{(\text{society})} = \frac{72{,}500€ - 80{,}100€}{92{,}200€ - 124{,}200€} = \frac{-7{,}600€}{-32{,}000€} = 0.2375 \qquad \textbf{Equation 17}$$

$\text{iCBR}_{(\text{society})}$ = incremental cost-*benefit ratio from society's perspective*

iCBR shows that for each 0.2375 euro spent on intensive care treatment, the society would earn 1 euro per additional survivor.

b.) net benefit (Equation 18):

$$\text{NB}_{(\text{society})} = (92{,}200€ - 124{,}200€) - (72{,}500€ - 80{,}100€) = \\ = 32{,}000€ - 7{,}600€ = 24{,}400€ \qquad \textbf{Equation 18}$$

$\text{NB}_{(\text{society})}$ = *net benefits from society's perspective*

The net benefits to the society from the intensive care treatment would be 24,400€ per additional survivor.

Exercises

Task 1

In table 5 are presented data for undertaking an economic analysis. Undertake and comment:

1. a CMA comparison of treatment B versus treatment A from: (a) the hospital's perspective and (b) society's perspective.
2. a CEA comparison of intensive care from society's perspective.
3. a CUA comparison of intensive care from society's perspective.
4. a CBA comparison of intensive care from society's perspective.

Table 5. Evaluation of treatment B versus treatment A

Cost or consequence	Treatment A	Treatment B
Cost per additional survivor (to hospital discharge)	2,200€	15,100€
Cost per additional survivor (to death)	60,500€	90,700€
Survival rate (to hospital discharge)	56.4%	91.1%
Survival time (per live birth):		
a. Life-years	22.2	30.8
b. QALYs	16.1	40.6
Earnings	88,300€	134,900€

Task 2

In four groups, perform an internet search of bibliographic data-bases (e.g. Medline, Pubmed) to find cases of four types of economic evaluation analysis, each group respectively. The groups should critically discuss the strengths and weaknesses of the selected cases and than present them to other students.

Task 3

The four groups have the roles of decision-makers. They are considering the approval of cases selected in Task 2. What additional information would they want to consider along with economic evaluation? Discuss in groups and present the findings to other students.

References

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of healthcare programme. Oxford: Oxford University Press; 2005.
2. Zerbe RO, Bellas AS. A primer for benefit-cost analysis. Northampton: Edward Elgar Publishing; 2006.
3. Schulman KA, Seils D. Clinical Economics. In: Interactive Textbook on Clinical Symptom Research. Available from URL: http://painconsortium.nih.gov/symptomresearch/bio_jlynn.htm. Accessed: May 30, 2013.

4. Briggs AH. Handling uncertainty in economic evaluation. In: Drummond MF, McGuire A, eds. Economic evaluation in health care: merging theory with practice. Oxford: Oxford University Press; 2001: p.172-214.
5. Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems. In: Bjegovic V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004: p.376-403. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: May 30, 2013
6. Rascati KL. Essentials of pharmacoeconomics. Philadelphia: Wolters Kluver/Lippincot, Williams&Wilkins; 2009.
7. Prieto L, Sacristan JA. Problems and solutions in calculating quality-adjusted life-years (QALYs). Health Qual Life Outcomes. 2003;1: 80.
8. Brent RJ. Cost-benefit analysis and health care evaluations. Northampton: Edward Elgar Publishing; 2003.
9. Phelps C, Mushlin A. On the near equivalence of cost-effectiveness and cost-benefit analyses. Int J Technol Assessment Health Care. 1991;7:12-21.
10. Birch S, Gafni A. Cost effectiveness/utility analyses. Do current decision rules lead us to where we want to be? J Health Econ. 1992;11:279-96.
11. Weinstein M, Zeckhauser R. Critical ratios and efficient allocation. J Pub Economics. 1973;2:147-158.
12. Weinstein M, Stason W. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977;296:716-721.
13. Robinson R. Cost-benefit analysis. Br Med J. 1993;307:859-862.
14. Byford S, Raftery J. Perspectives in economics evaluation. Br Med J. 1998;316:1529.
15. Coyle D, Drummond MF. Costs of conventional radical radiotherapy versus continuous hyperfractionated accelerated radiotherapy (CHART) in the treatment of patients with head and neck cancer and carcinoma of the bronchus. Centre for Health Economics, University of York, York; 1997.

Recommended readings

1. Brent RJ. Cost-benefit analysis and health care evaluations. Northampton: Edward Elgar Publishing; 2003.
2. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.
3. Rascati KL. Essentials of pharmacoeconomics. Philadelphia: Wolters Kluver/Lippincot, Williams&Wilkins; 2009.
4. Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems. In: Bjegovic V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company, 2004: p.376-403. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: May 30, 2013.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Strengthening public health services: evaluation of essential public health operations**
Module: 2.38	ECTS (suggested): 0.5
Author(s), degrees, institution(s)	Fimka Tozija, Prof. MD, PhD, Institute of Public Health, Faculty of Medicine, Skopje, Republic of Macedonia Dragan Gjorgjev, Prof. MD, PhD, Institute of Public Health, Faculty of Medicine, Skopje, Republic of Macedonia Dance Gudeva Nikovska, MD, MPH, Teaching Assistant, Faculty of Medicine, Skopje, Republic of Macedonia
Address for correspondence	Fimka Tozija, Prof. MD, PhD, Institute of Public Health 50 Divizija No 6 1000 Skopje, Republic of Macedonia Tel: +38923125044/110 E-mail: ftozija@mt.net.mk
Keywords	Essential Public Health Functions, Essential Public Health Operations, evaluation, Public health services, tool.
Learning objectives	After completing the module students should become familiar with the ten essential public health operations, become familiar with the tool and be able to apply the tool in their countries
Abstract	This module is consistently based on a clear statement concerning public health and health systems, including definitions, boundaries and concepts, in view of differences in the ways in which European health systems and public health services are organized, operated and governed. The module presents a revised set of ten horizontal essential public health operations (EPHOs), including the core public health services within each one of them, to become the unifying and guiding basis for any European health authority to set up, monitor and evaluate policies, strategies and actions for reforms and improvement in public health. Repeated measurements over time facilitate consistent quantification between measurement and identification of the "grey zones" in public health system, thus facilitating design of targeted interventions for institutional capacity strengthening.
Teaching methods	Qualitative method, SWOT analysis, instrument application, individual work, group work.
Specific recommendations for teachers	To be familiar with the instrument, the health system and the needs of the group.
Assessment of students	Formative assessment - during the course, individual and group work. Summative assessment – final SWOT analysis applying the instrument.

STRENGHTENING PUBLIC HEALTH SERVICES: EVALUATION OF ESSENTIAL PUBLIC HEALTH OPERATIONS

Fimka Tozija, Dragan Gjorgjev, Dance Gudeva Nikovska

Scope of public health services (PHS) and activities

Essential public health functions (EPHFs)

Public health capacities and services are underpinned by the Acheson definition of public health (1): *"Public health is the science and art of preventing di*sease, prolonging life and promoting health through *the organized efforts of society."*

A unifying principle of public health is its essentially "public" nature and the fact that it is mainly focused on the health of the whole population. Public health can be understood as a key aspect of the wider health system and can play an important role in improving the effectiveness and efficiency of health system delivery.

The definition of a health system adopted in the Tallinn Charter in 2008 (2) is retained: "Within the political and institutional framework of each country, a health system is the ensemble of all public and private organizations, institutions and resources mandated to improve or restore health. Health systems encompass both personal and population services, as well as activities to influence the policies and actions of other sectors to address the social, environmental and economic determinants of health."

'Essential' is understood as being fundamental and even indispensable to meeting public health goals and to defining public health. This term also makes reference to the definition of the responsibilities of the state, through health authorities, considered essential to the development and practice of public health. Consequently, the essential operations are at the core of the definition of the entire public health field and, in turn, are indispensable to improving the health of populations. Other activities may or may not be added, but the essential ones form the core of public health activity.

The approach of identifying essential areas of public health activity and services was an integral component in the elaboration of the WHO's Health for All (HFA) policy. The WHO has been involved in the ongoing discussion around defining the scope of public health and its essential services, including a series of worldwide discussions in 1997, involving 145 prominent leaders of public health in 67 nations from all regions of the world.

The international Delphi study aimed to describe "essential public health functions" has defined the concept of "essential public health functions", tried to establish a consensus on what activities and services are essential and to confirm which public health activities and services require the elaboration of performance standards (3). What is the minimum absolutely necessary to protect the health of the population? Work has also been undertaken at national level, in, for example, the United States (4) and the United Kingdom (5).

Modified "essential public health functions" are also published in the relevant publications of the Pan American Health Organization and WHO Regional Office for Western Pacific (6). In this publication EPHF were defined as: a set of fundamental activities that address the determinants of health, protect a population's health, and treat diseases (of public health significance). These public health functions represent public goods, and in this respect governments would need to ensure the provision of these essential functions, but would not necessarily have to implement and finance them. They prevent and manage the major contributors to the burden of disease by using effective technical, legislative, administrative, and behaviour-modifying interventions or deterrents, and thereby provide an approach for intersectoral action for health.

This approach stresses the importance of numerous different public health partners. Moreover, the need for flexible, competent state institutions to oversee these cost-effective initiatives suggests that the institutional capacity of states must be reinforced (7,8).

Public health experts now believe that public health agencies should perform a set of ten defined essential public health functions that new developments on international arena of public health confirm (9). WHO has described the components of a functioning health system as including (among others) six essential building blocks: good health services; a well-performing health workforce; a well functioning health information system; equitable access to medical products, vaccines and technologies; a good health financing system that raises sufficient funds for health and assures access; and leadership and governance (10). WHO has presented the latest developments regarding ten essential public health operations at the RC61 in 2011 (11).

The evaluation of PHS in Europe has built on the work in the context of the particular challenges parts of Europe face in undertaking reform of their public health infrastructures, as well as the usefulness of analyzing PHS and activities within the framework of the four health systems functions of Stewardship, Resource Generation, Finance and Service Delivery (12).

For reasons of clarity in discussing the interaction between the health system functions and the "essential public health functions", the components of public health are called "operations". As such, the evaluation of PHS in this Module is based on a set of essential public health operations, which elaborate the details of the health system functions in the field of public health (13).

It is worth noting that these essential public health functions vary according to organization. While most definitions have much in common, the Centres for Disease Control and Prevention (CDC), the Pan American Health Organization (PAHO) and WHO all have slightly different ideas of what constitute essential public health functions (13).

Methodology of the assessment process

The methodology of Essential Public Health Functions (EPHF) analysis is explorative in nature. The approach is analogous to a "polls of experts", emphasizing that "for many dimensions of governance, subjective data is the only data that is potentially informative". Survey performed with this instrument is qualitative and conclusions have to be cautiously interpreted due to limitations in its design and methodology. Essentially, the methodology is designed to obtain self-evaluation of the public health system by people who know the system well from within (14), rather then providing for "objective" information typically asked in surveys. Results obtained with this kind of survey should facilitate provision of recommendations for improvement of the essential public health functions, reduction of barriers and costs associated with them, identification of critical issues and dilemmas, as well as development of strategies to address them.

Application of this methodology in USA and South American countries have shown that despite the fact that scoring criteria are not completely defined, the instrument is an excellent tool for identification of strengths and weaknesses of public health systems. Repeated measurements over time facilitate consistency quantification between measurement and identification of the "grey zones" in public health system, thus assisting in design of targeted interventions for institutional capacity strengthening (6).

Measurement instrument

The basic instrument for this kind of assessment is a standardized questionnaire designed by US CDC and PAHO. In USA, the instrument has been used by several countries for identification of most striking problems in public health services, and consequently used for analysis of the health systems in Latin American countries (6).

The questionnaire is divided into 11 segments labeled Essential Public Health Functions (EPHF) - Table 1. Each function is assessed through set of predefined indicators (total of 49 indicators), specific for each segment. Every segment begins with description of services covered by the respective function; segments are divided into sub-segments with questions specified for the defined indicators.

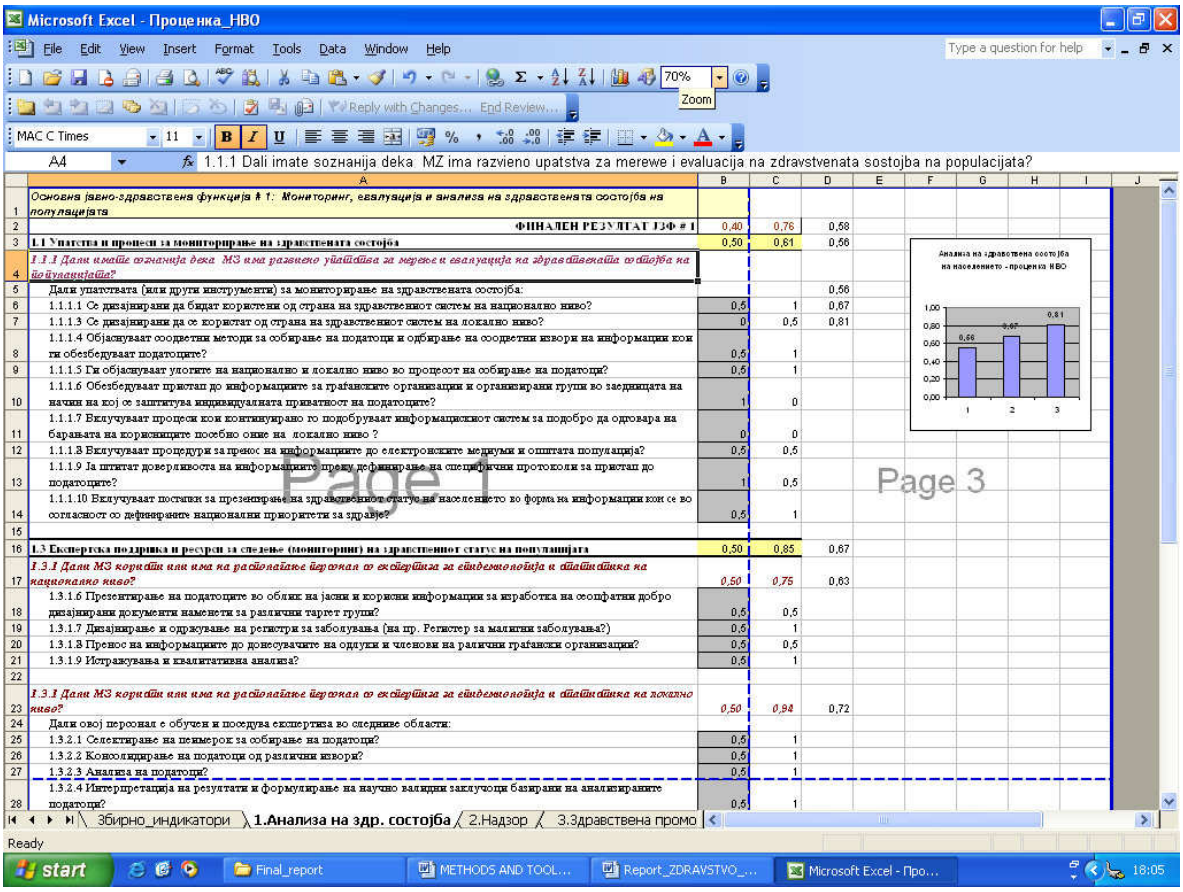
Table 1. Eleven essential public health functions

ESSENTIAL PUBLIC HEALTH FUNCTIONS
1. Monitoring, evaluation and analysis of health status
2. Public health surveillance, research and control of risks and threats to public health
3. Health promotion
4. Social participation in health
5. Development of policies and institutional capacity for planning and management in public health
6. Strengthening of institutional capacity for regulation and enforcement in public health
7. Evaluation and promotion of equitable access to necessary health services
8. Human resources development and training in public health
9. Ensuring the quality of personal and population-based health services
10. Research in public health
11. Reducing the impact of emergencies and disasters on health

Source: PAHO, 2002 (5).

The instrument is simple, user friendly Excel file. Separate sheet is assigned to each individual EPHF, in addition to the summary sheet for all 11 functions. Graphical presentation of sub-segments is automatically displayed, as scores for each question are entered into designated cells. Individual answers are recorded as YES (response is assigned value of 1) or NO (response is assigned value of 0). Upon completion of each segment, summary value is obtained for individual indicators (result between 0 and 1, depending on proportion of positive answers for that indicator). Mean value is calculated from the results obtained for each indicator by individual respondents, in order to assess the respective EPHF (Example of the excel sheet for assessment of one EPHF is presented in Figure 1).

Figure 1. Example of the Excel file for assessment of one EPHF (13)



Application of the instrument

Assessment of public health system functioning in a given country could be performed in two different ways:

- by a representative sample of individual respondents from relevant agencies. In this case, the instrument is administered individually and each score represents individual opinion of the person to each of the 11 EPHF. Mean value is then calculated for the respective indicator, as explained above, or
- at a meeting or workshop, convening professionals from relevant agencies (health personnel, academicians, civil sector representatives, other relevant specialists). At the end of the workshop, results of the measurement are usually shared with participants, providing examples of the types of analysis that can be done when measuring EPHF, oriented toward the identification of intervention areas in order to improve the institutional capacity of health authorities in exercising EPHF that relate to it. Number of members of each group is arbitrary; however, experience shows that it ranges from 15-30 participants representatives of relevant Governmental and Non-Governmental institutions.

Scoring/results of the measurement

The score for each indicator for each EPHF is based on the scores obtained from so-called First Tier Statements. Individual questions for each EPHF (measures and sub-measures pertaining to that statement), could be answered YES (response is assigned value of 1), NO (response assigned with 0) or partial response (0.5). The instrument can be modified with an option to assign 0.5 for answers that could not be answered simply YES or NO; examples include answers such as "partial implementation, existing, but not enforced law/regulation, in phase of preparation etc." The calculation of the final score of every first tier statement is essentially the average of the "Yes" and "Partial" responses to the measures and sub-measures. The score assigned to the indicator is the average of results obtained for each of the measures within the indicator and the average of the results of all the indicators in a function determines the score for the performance of that particular EPHF.

Interpretation of results

Conventional scale is proposed for overall interpretation of obtained results of EPHF analysis. It is presented in Box 1.

Box 1. Conventional scale for overall interpretation of results of essential public health functions analysis

76-100% (0.76-1.0)	Quartile of **optimal** performance
51-75% (0.51-0.75)	Quartile of **above average** performance
26-50% (0.26 - 0.5)	Quartile of **below average** performance
0-25% (0.0 - 0.25)	Quartile of **minimum** performance

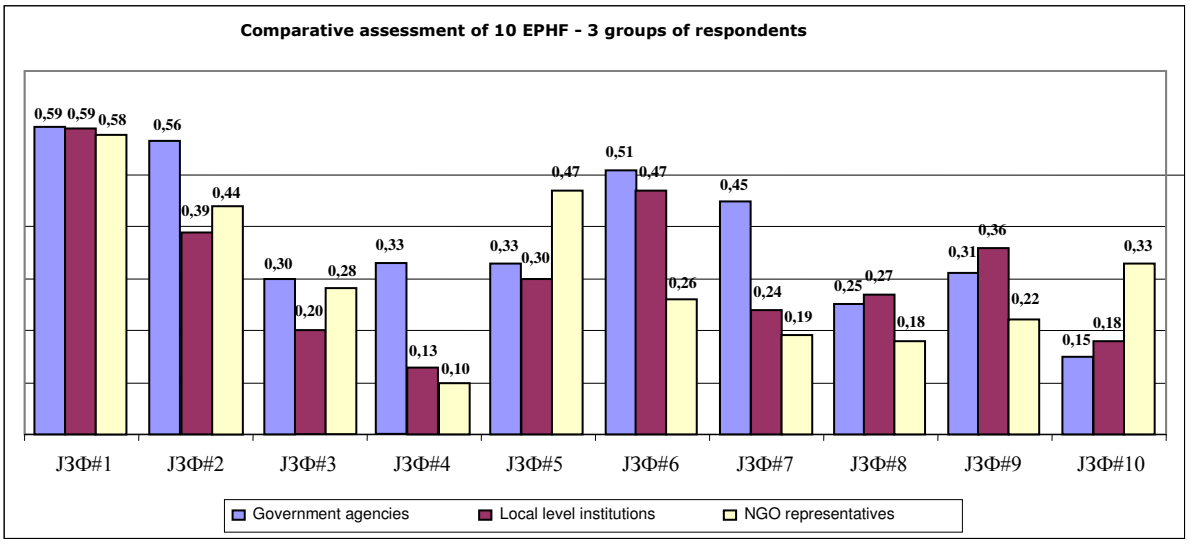
Box 2: Evaluation of EPHFs in Republic of Macedonia

Methods: The analysis has been performed in the period Apr-Nov, 2007 using the standardized questionnaire developed by US CDC and PAHO (14) translated and adapted to Macedonian context. The instrument incorporates 11 areas of EPHF and outlines 49 indicators that are evaluated through answers on questions for measures and sub measures in the defined area. It was agreed to exclude EPHF#11 Reducing the impact of disasters and emergencies to health, in order to provide more detailed analysis of the other 10 EPHF. The survey was performed in three phases. Interviews were conducted with 3 groups of respondents – central government officials, representatives of government institutions at local level and representatives of non-governmental sector.

Results: Comparative analysis of all 10 EPHF, applying conventional interpretation of results, shows that none of the EPHF prove optimal results, a score shared by all three groups of respondents. Highest score is recorded for EPHF#1 Monitoring, evaluation and health situation analysis, while lowest score is documented for EPHF#3 Health promotion, EPHF#8 Human Resource Development and Training in Public Health and EPHF#10 Research in Public Health. The summary results of EPHF analysis are presented in Figure 2.

Assessment of the specific indicators for each EPHF has identified existence of "grey zones", which were used to perform subsequent analysis of strengths, weaknesses, opportunities, and threats (SWOT analysis) of governance in public health sector. Recommendations were given for remedial activities for each EPHF.

Figure 2. Assessment of 10 essential public health functions by 3 groups of respondents



Source: Tozija F., Gjorgjev D., Gudeva Nikovska D. 2010 (13).

The Essential Public Health Operations (EPHOs) in Europe

The operations represent all the activities within the area of public health that are essential to the delivery of adequate and modern PHS. They cut across the health system functions to some degree but are concentrated in service delivery. They are divided into two groups: those focused on recognizable services, and those that define, support, feed into and enable those services to be delivered – in effect instrumental operations. These latter dimensions deal mainly with areas of stewardship, but also with resources (2).

The essential public health operations comprise the core public health areas of practice, which include: strategy development, workforce development, quality assessment, health information, health promotion, communicable disease surveillance, chronic disease prevention, public health dentistry, environmental health, occupational safety and health, injury prevention, food safety, public health nutrition, mother and child health, community genetics, global health and public health laboratories, and others (15).

The services and activities that make up the essential operations are usually not supplied by a single public health institution, body or profession. Indeed one benefit of the approach is precisely to identify

horizontal activities rather than focus on the activities of institutions, in order not only to assess whether essential operations are being carried out but also so the approach can be applied to different institutional settings. Many PHS are to be practiced and delivered by many of the public health areas of practice at the same time. This may be through structures, institutions or groups of professionals within particular areas of public health, such as communicable disease surveillance, occupational health, environmental health or food safety. The dimensions are specifically intended to be applicable to all health system structures in the area of public health, across all countries, so that they are useful regardless of the type or particular structure of a health system (16).

Methodology of the assessment process

The evaluation of PHS undertaken with the WHO Regional Office for Europe, is based on a defined set of "essential public health operations", which have their origins in attempts to identify the most important and fundamental aspects of PHS and activities.

The EPHOs, of which there are 10 at present, are a work in progress and are currently being used in a self assessment programme in the European Region (17). The 10 Essential Public Health Operations (EPHOs) form the framework for a practical web-based tool developed by the WHO Regional Office for Europe to help countries assess their public health services and activities (18).

Measurement instrument

The self-assessment tool for the evaluation of public health services has been piloted and is currently being updated. The self-assessment tool can be used: (1) to create, through the definition of essential operations, a common understanding of essential public health services and activities; (2) to inspire discussion and debate about essential services and activities in the field of public health within ministries of health, institutes of public health and among other public health experts and actors; (3) to provide a framework for such experts to assess their own country's performance across these services and activities; and (4) to produce a relatively quick analysis of the principal strengths and weaknesses in public health services and activities as a basis for devising strategies, plans, reforms and further evaluation of public health services, in general or in particular areas.

Tool has been used by key experts from all areas of public health in various countries to identify strengths and weaknesses and areas in need of investment and reform. Report on public health capacities and services in the European Region (19) provides an overview of the current status of public health services and capacities across the WHO European Region, with a view to strengthening the development of future public health services and capacities. It aims to underpin and complement the European Action Plan for Strengthening Public Health Capacities and Services (EAP) (20). The EAP has been drawn up in response to the WHO resolution EUR/RC61/R2, based on a broad consultative process and is a key pillar of the overarching regional policy framework, Health 2020 (21).

Application of the instrument

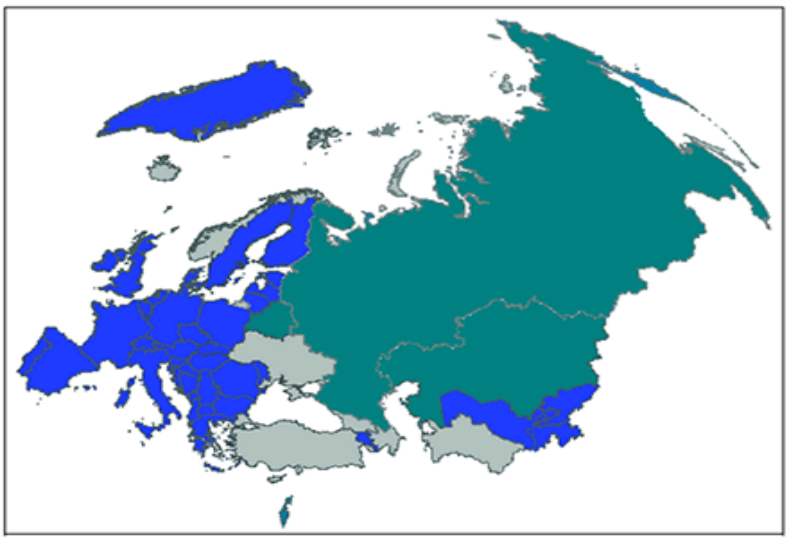
The evalution report is part of a series of three studies being conducted by the WHO Regional Office for Europe. These are a review of policy and legislation instruments and tools for public health; a "snapshot" review of organizational models for delivering essential public health operations (EPHOs) and public health services; and a summary of country assessments of public health capacities and services. These studies support the development of the EAP.

The information is derived from assessments of public health services in 41 of the 53 Member States in the WHO European Region and it includes assessments from an EU study and the self-assessment tool. Figure 3 shows in blue those countries that have completed the EU or WHO self-assessment. Countries with assessments in progress are marked in green (Belarus, Israel, Kazakhstan and the Russian Federation). Those highlighted in grey (Azerbaijan, Georgia, Iceland, Norway, San Marino, Turkey, Turkmenistan and Ukraine) have not yet undertaken a self-assessment and do not plan to do so imminently.

Evaluation consists of self-assessment reports from 17 countries performed using the European Region self assessment tool. It also includes findings from a review of public health capacities in the 27 European Union (EU) countries – a study for the European Commission by Maastricht University (Maastricht University, unpublished). In addition, two countries also conducted an assessment with the European Observatory on Health Systems and Policies. Some countries have conducted more than one type of assessment. Azerbaijan, Georgia, Iceland, Norway, San Marino, Turkey, Turkmenistan and Ukraine are yet to commence any assessment.

Finally, other relevant WHO work looking at existing activity related to the 10 essential public health operations (EPHOs) that form the basis of the EAP was included in the report, based on the WHO strategic objectives for 2008–2013. These EPHOs were updated during the consultation process to include a new area on advocacy, communication and social mobilization (EPHO 9), which was not fully captured during the assessments.

Figure 3. Self-assessment of public health services in Member States in the WHO European Region



Source: Review of public health capacities and services in the European Region. WHO 2012 (19).

Results of the measurement

The report presents a summary of the main findings by EPHO (19).

Box 3. The 10 Essential Public Health Operations (EPHOs) 2012

1. Surveillance of population health and well-being
2. Monitoring and response to health hazards and emergencies
3. Health protection, including environmental, occupational, food safety and others
4. Health promotion, including action to address social determinants and health inequity
5. Disease prevention, including early detection of illness
6. Assuring governance for health and well-being
7. Assuring a sufficient and competent public health workforce
8. Assuring sustainable organizational structures and financing
9. Advocacy, communication and social mobilization for health
10. Advancing public health research to inform policy and practice

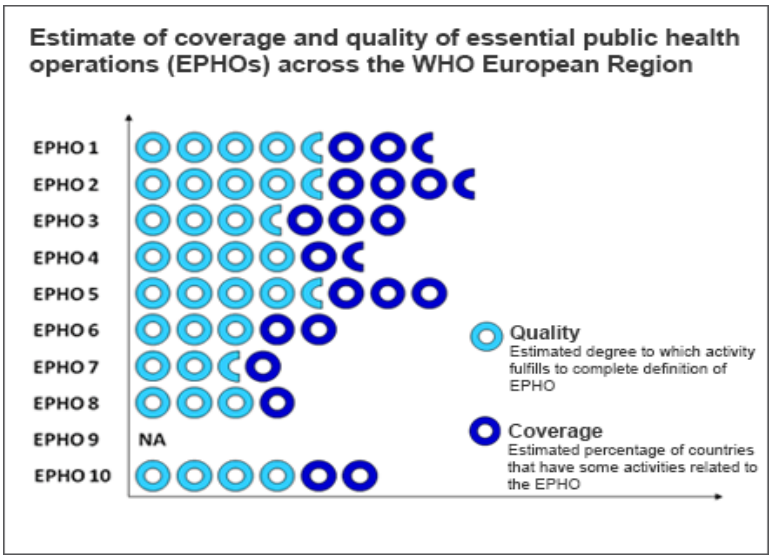
Source: Review of public health capacities and services in the European Region. WHO 2012.

The main findings across the EPHOs are summarized below (19).

- Across the Region, the strongest geographical coverage and quality is for EPHOs 1–3, including surveillance, monitoring, emergency planning, immunization, environmental health and health protection.
- The less well developed EPHOs include EPHO 4 on health promotion, inequalities and he wider determinants of health; surveillance to address NCDs is also weak – this pattern s found especially in the CIS countries.
- The enabling EPHOs 6, 7, 8 and 9 are also less well developed across the Region, addressing governance, workforce development, financing and communications – these are generally weaker in the CIS countries.
- Where there are greater health inequalities there are generally less well developed public health services and capacities, illustrating the inverse care law in an approximate line from north-west to south-east across the Region, with central Asian countries experiencing greatest health inequalities and least capacity to address them.
- The main public health challenges facing the Region need core EPHOs 1–5 particularly to be strengthened; additionally, governance and communication (EPHOs 6 and 9) are considered highly relevant.
- All the EPHOs were found to be relevant to a greater or lesser e tent for WHO strategic objectives and categories, contributing to overall strengthening of WHO work and illustrate the need to take an integrated, horizontal approach to delivering public health services.

These findings are illustrated in the Figure 4 below:

Figure 4. Estimate of coverage and quality of EPHOs across the WHO European Region

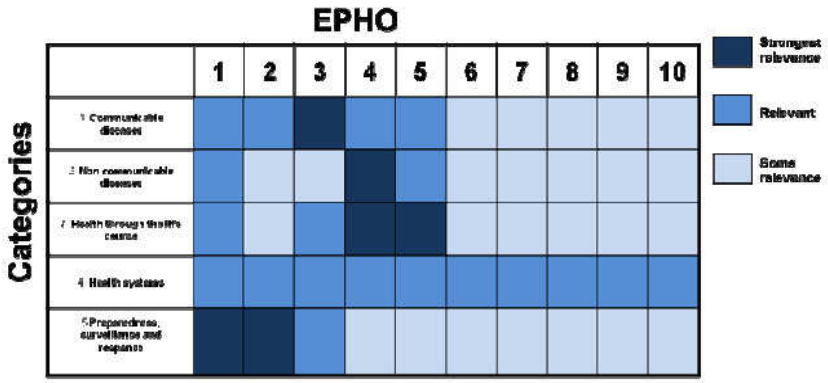


Source: Review of public health capacities and services in the European Region. WHO (19).

Across the Region, the strongest public health responses are for surveillance, monitoring, emergency planning, immunization, environmental health and health protection. Weaker areas include health promotion and action to address inequalities and the wider determinants of health. NCD surveillance is also weak. Governance, workforce development, financing and communications are also less well developed across the Region; this pattern is found especially in the Newly Independent States (NIS) (19).

In the future, WHO will be changing from 11 strategic objectives to 5 main categories, for which the degree of relevance for each of the EPHOs is also estimated and marked in Figure 5. Significantly, all categories have at least some relevance to the EPHOs, with greatest relevance to the core EPHOs (1-5). However, the enabling EPHOs (6-10) are still very important in strengthening overall services. The "health systems" category includes public health services, and is therefore marked as relevant across all EPHOs.

Figure 5. Relationship between the WHO categories and the EPHOs Recommendations



Source: Review of public health capacities and services in the European Region. WHO 2012 (19).

Discussion – evidence based policy interventions

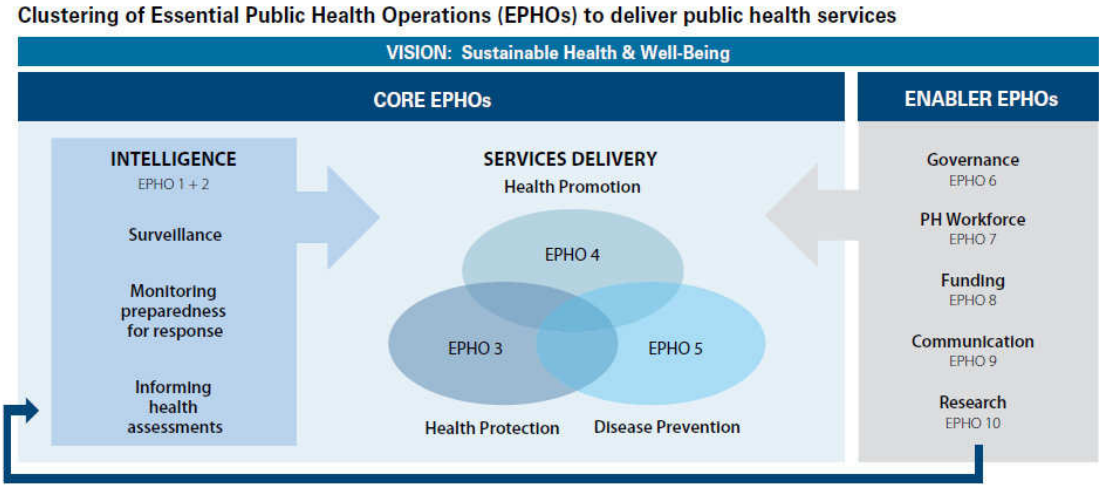
The evidence gathered through the process of the WHO/Europe-led PHS evaluation in over 20 European countries, the Review of Public Health Capacities in the European Union and additional studies on policy tools and instruments for public health, as well as a "snapshot analysis" of organizational models for delivering the EPHOs, outlines some of the major challenges facing health policies and systems in the WHO European Region, including consideration of public health services and infrastructures and the public health aspects of health care services (19).

Member States and the Regional Office and its international partners intend to follow in order to strengthen public health capacities and services and secure the delivery of the ten EPHOs in an equitable way across the whole Region (22).

The most effective and efficient method to deliver these operations is through an integrated approach, rather than in vertical programmes, so the EPHOs have been clustered into two groups of five: core EPHOs (requiring public health skills and expertise to deliver them) and enabler EPHOs. There is an additional

cluster of public health intelligence EPHOs (1, 2 and 10) which are based on monitoring, surveillance and research. The main focus of service delivery – health promotion, health protection, and disease prevention (EPHOs 3, 4 and 5) is via public health services. The enablers (EPHOs 6–10) include strengthening governance, workforce development, financing, communication and research (20) (Figure 6).

Figure 6. Clustering EPHOs to deliver public health services



Source: European Action Plan for Strengthening Public Health Capacities and Services. WHO. 2012 (20).

Member States collaborated actively in the development and strengthening of the EAP, and in reviewing and strengthening the EPHOs to ensure full consistency with Health 2020, particularly in relation to a "whole-of-government" approach to improving health, acting on the structural and social determinants of health and tackling health inequalities. Key areas for action are addressed in relation to further developing, strengthening and sustaining existing public health capacities and services, with the aims of improving health and tackling health inequalities through action on the social determinants of health (22).

Key messages:

- Strengthen the delivery of public health services by developing and integrating health promotion and disease prevention with robust health protection services.
- To support service delivery, the enablers for public health that especially need further development include governance, workforce development, financing and communication.
- Focus public health services to ensure they address inequalities and the wider determinants of health to achieve the overall vision of promoting health and well-being in a sustainable way.

Case study

The SEE Health Network and the evaluation of public health services

The Evaluation of Public Health Services in SEE countries and the production of National and Regional reports on PHS is a key undertaking of the South-eastern European Health Network (SEEHN) (23). The evaluation has formed the first component of the PHS project "Strengthening public health capacity and services".

The PHS evaluation marks a major step forward in understanding the challenges faced by the nine countries in reforming and developing their PHS, as well as the opportunities available to them to improve the health of their populations. The National and Regional reports represent a valuable addition to the sum of knowledge about the state of PHS in the SEE countries, to the sharing of information and experience between them, as well as a platform on which to build appropriate policies, infrastructures and services for the development of PHS in the region.

Methodology

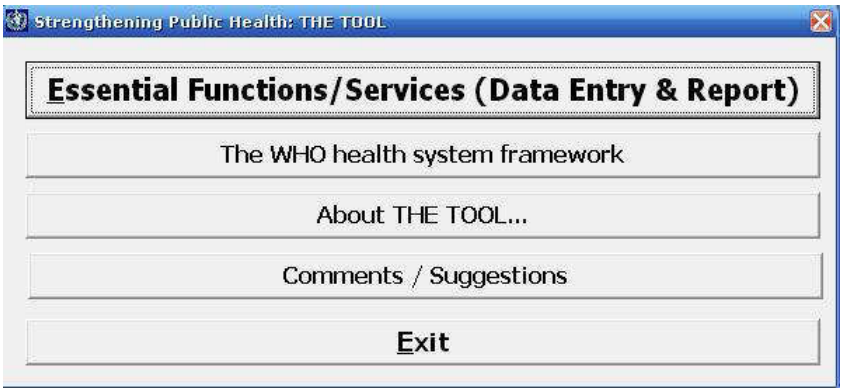
The evaluation process was undertaken by National Focal Points in each of the countries and managed and coordinated by a Regional Project Manager. The focal points were responsible for managing the whole process at national level, gathering information and bringing together the necessary expert opinion to make assessments of services. The Regional Project Manager was responsible for process management and the writing of the regional report, supported by WHO.

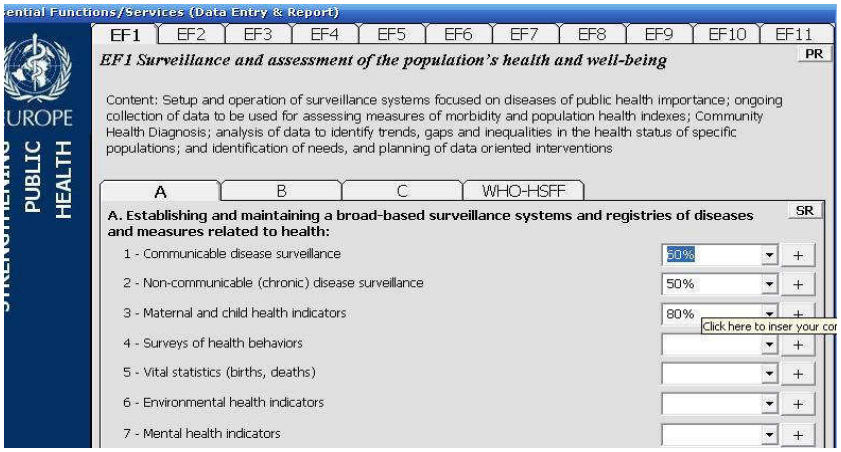
Each National Focal Point was responsible for managing the process of evaluation. To gather information, key actors within the national public health institutions were collectively responsible for

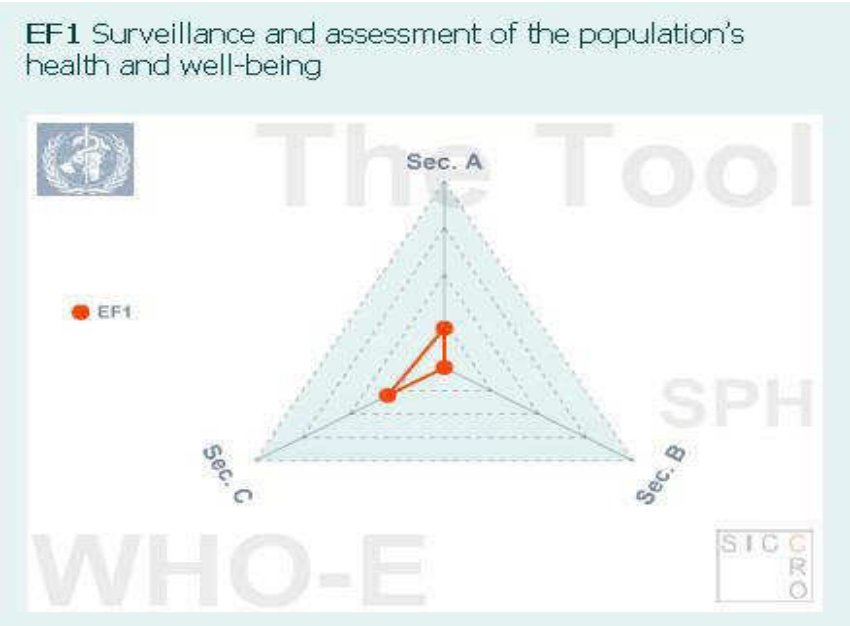
answering a comprehensive questionnaire. Expert seminars and workshops were used in some cases to agree the main findings national recommendations and conclusions. Regional and international experts were also convened in workshops during the process to discuss information and findings and to contribute to the regional-level analysis of strengths and weaknesses and agreed recommendations.

Measurement instrument

The tool has been conceived as a guiding instrument (to ensure that all essential functions and services are taken into account) and also useful for self-assessment.







The tool is conceived as initially generic, but its design should also be flexible/adjustable to the specific context and needs of each Member State of WHO/Euro Region. Language adjusted and context sensitive

Results

The tool gives an overview of the whole range of public health services, within the health system enabling their delivery in a strategic and integrated way, enabling each Member State to realize the gap between the current situation of the Public Health Services and the expected one.

Swot analysis

Strengths	Weaknesses
• Established public health institutes and legal frameworks for public health • Intersectoral arrangements for public health programmes • Clear and legally defined package of public health services • Population health surveillance systems well established • Capacity for epidemiological research • Supportive international framework with many charters and agreements in the area of public health	• Lack of integration of policy across sectors, including intersectoral assessment of the results of policies, strategies and programmes • Lack of integration of social determinants of health in policies • Lack of data disaggregated by socioeconomic groups • Weak capacity for monitoring policy implementation • Lack of quality assessment for health care services, requiring development of standards • Lack of monitoring of private activity
• **Opportunities**	• **Threats**
• Opportunity for devolved decision-making, including local accountability for services • Coordination of institutions, agencies and ministries competent in public health services • Establishment of processes for evaluation and monitoring of cost–effectiveness of public health programmes and activities • Existing good practice in intersectoral cooperation can be built on and extended; multisectoral cooperation for the broader health determinants can be established • Public health agenda in EU, and harmonization of the regulatory framework; the EU is a driver of positive change and the accession process a significant opportunity • Accreditation of laboratory resources to identify and close gaps – opportunity for regional cooperation	• Over decentralization may present risks of increased differences in the health conditions and provision of services among the different parts of the country • Privatization process without a strategic approach could worsen problems in the distribution of resources • Threat to economic effectiveness of public health due to lack of intersectoral arrangements • Poverty is an underlying threat to population health and to effectiveness of health system • Any disengagement from EU accession process affecting strategy or other factors such as economic growth, and therefore funding • Insufficient mechanism for maintaining public health workforce standards • Low salaries in the public health professions might undermine recruitment and lead to potentially severe shortages of personnel in some areas

Source: Evaluation of public health services in south-eastern Europe. WHO, 2009 (23).

Exercises

Task 1: Students will split into three groups. Each group will:

- analyze and compare SWOT analysis for the EPHOs for EU, SEE and HIS;
- group presentations (outlined in a handout).

Task 2: Students will split into three groups. Each group will:

- analyze all EPHOs: findings and recommendations;
- group presentations – overcoming the identified gaps for each EPHO.

References

1. Public health in England. The report of the Committee of Inquiry into the Future Development of the Public Health Function. London, HMSO, 1988.
2. The Tallinn Charter: Health Systems for Health and Wealth. Copenhagen, WHO Regional Office for Europe, 2008, available from

http://www.euro.who.int/__data/assets/pdf_file/0008/88613/E91438.pdf.

3. Bettcher DW, Sapirie S, Goon EH. Essential public health functions: results of the international Delphi study. World Health Stat Q. 1998;51:44–54.
4. U.S. Department of Health and Human Services. Public Health Service. For a Healthy Nation: Returns on Investment in Public Health. Washington, D.C. 1995.
5. Faculty of Public Health Medicine. Strengthening Public Health Function in the UAE and EMR of the WHO. London: Faculty of Public Health Medicine of the Royal Colleges of Physicians of the United Kingdom. 2001.
6. Pan American Health Organization. Public Health in the Americas. Conceptual Renewal, Performance Assessment, and Bases for Action. Washington, DC: OPS, available from http://www.paho.org/english/hia_1998ed.htm.
7. Addressing key public health and health policy challenges in Europe: moving forwards in the quest for better health in the WHO European Region. Copenhagen, World Health Organization, 2010 (resolution EUR/RC60/R5), available from http://www.euro.who.int/data/assets/pdf_file/0018/122229/RC60_eRes5.pdf. Accessed 11 September 2013.
8. Yach D. Redefining the scope of public health beyond the year 2000. Current Issues in Public Health, 1996,2:247−252.
9. Concept note for developing a public health strategy for Europe. Copenhagen, World Health Organization, 2010 (document EUR/RC60/SCRC/18) ovoj document go nema, se spomnuva na ovoj link http://www.euro.who.int/__data/assets/pdf_file/0009/148266/RC61_einfdoc01.pdf .
10. Everyone's business: strengthening health systems to improve health outcomes: WHO's framework for action. Geneva, World Health Organization, 2007, available from http://www.who.int/healthsystems/strategy/everybodys_business.pdf.
11. WHO Resolution on strengthening public health services, available from http://www.euro.who.int/__data/assets/pdf_file/0006/173616/RC62rs05-EAP-ENG.pdf.
12. WHO World Health Report 2000, available from http://www.who.int/whr/2000/en/.
13. Tozija F., Gjorgjev D. Gudeva Nikosvka D. Measurement, monitoring and evaluation of public health systems: Assessment of essential public health functions. In: Editors: Kragelj Lj, Bozikov J. Methods and Tools in Public Health. Hans Jacobs Verlag. 2010:831-847.
14. The Essential Public Health Functions as a Strategy for Improving Overall Health Systems Performance: Trends and Challenges since the Public Health in the Americas Initiative, 2000-2007.
15. Developing a framework for action for strengthening public health capacities and services in Europe – Interim draft. Copenhagen, WHO Regional Office for Europe, 2011, available from http://www.euro.who.int/__data/assets/pdf_file/0009/148266/RC61_einfdoc01.pdf.
16. Developing a framework for action for strengthening public health capacities and services in Europe – Interim draft. Copenhagen, WHO Regional Office for Europe, 2011, available from http://www.euro.who.int/__data/assets/pdf_file/0009/148266/RC61_einfdoc01.pdf.
17. Commission on the Social Determinants of Health. Closing the gap in a generation: health equity through action on the social determinants of health. Geneva, World Health Organization, 2008, available from http://www.who.int/social_determinants/thecommission/finalreport/en/index.html.
18. WHO Health Systems. 2013, available from http://www.euro.who.int/en/what-we-do/health-topics/Health systems/public-health-services/activities/assessment. Accessed on 15 March 2013.
19. Review of public health capacities and services in the European Region. WHO 2012, available from:http://www.euro.who.int/en/what-we-do/health-topics/Health-systems/public-health services/publications2/2012/review-of-public-health-capacities-and-services-in-the-european-region.
20. European Action Plan for Strengthening Public Health Capacities and Services. WHO. 2012, available from http://www.euro.who.int/__data/assets/pdf_file/0005/171770/RC62wd12rev1-Eng.pdf.
21. Health 2020, available from http://www.euro.who.int/en/who-we-are/governance/regional-committee-foreurope/sixty-second-session/documentation/working-documents/eurrc629-health-2020-a-european-policyframework-supporting-action-across-government-and-society-for-health-and-well-being.
22. Strengthening services and capacity: an action plan for Europe. Promoting health and well-being now and for future generations. WHO. 2012, available from.

http://www.euro.who.int/__data/assets/pdf_file/0007/172762/Strengthening-public-health-services-and-capacity an-action-plan-for-Europe-Eng.pdf.

23. Evaluation of public health services in south-eastern Europe. WHO, 2009, available from http://www.coebank.org/Upload/infocentre/brochure/en/who_evaluation_of_public_health_services_in_south eastern_europe.pdf.

Recommended readings

1. World Health Organization, Pan American Health Organization. Public Health in the Americas. Conceptual Renewal, Performance Assessment, and Bases for Action. Washington, D.C.: Pan American Health Organization, 2002. Available from URL.
2. European Action Plan for Strengthening Public Health Capacities and Services. WHO. 2012.
3. Strengthening services and capacity: an action plan for Europe. Promoting health and well-being now and for future generations. WHO.
4. Evaluation of public health services in south-eastern Europe. WHO. 2009.
5. Review of public health capacities and services in the European Region. WHO 2012
6. Self-assessment tool for the evaluation of public health services [web site]. Copenhagen, WHO Regional Office for Europe. http://www.euro.who.int/en/what-wedo/health-topics/Health-systems/public-health-services/activities/self-assessment-tool-for-theevaluation-of-public-health-services.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Public health capacity building: adult education**
Module: 2.39	ECTS (suggested): 0.4
Author(s), degrees, institution(s)	Gordana Pavlekovic, MD, PhD, Associate Professor Andrija Stampar School of Public Health, Medical School, University of Zagreb, Croatia Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Anja Kragelj, Student of Andragogy, Chair of Andragogy, Faculty of Arts, University of Ljubljana, Slovenia
Address for correspondence	Gordana Pavlekovic Andrija Stampar School of Public Health, Rockefellerova 4, 10000 Zagreb, Croatia E-mail: gpavleko@snz.hr
Keywords	Adult education, andragogy, didactical methods.
Learning objectives	After completing this module students should: • know the definition and characteristics of adult education; • be familiar with different teaching methods used in adult education; • be able to use in practice some of group teaching methods of adult education.
Abstract	Like it holds in medicine that children are not diminished adults, and that paediatrics is not only curtailed internal medicine, it holds in educational process that adults are not augmented children. Adult education methods are in fact the same methods we could find in educational process of children, but the application of these methods is rather different. The module is describing principles of adult education, adult education process characteristics and teaching methods appropriate in public health capacity building, and provides some skills in using group education methods.
Teaching methods	An introductory lecture gives the students a first insight into the characteristics of adult education process, and understanding of phenomena. Afterwards, students first carefully read the recommended readings. Afterwards they form small groups. Every group is given a problem concerning some public health issue to be taught. In a group, students first use the proposed teaching method, and present the use of this method in front of other students. In continuation, then try to find the best teaching method for another public health theme. Every group prepares presentation of method, chosen as optimal. Other students critically discuss their choice.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple-choice questionnaire for individual assessment (50% of the score); assessment of group work (50% of the score).

PUBLIC HEALTH CAPACITY BUILDING: ADULT EDUCATION

Gordana Pavlekovic, Lijana Zaletel-Kragelj, Anja Kragelj

Theoretical background

Introduction

Teaching adults is not the same as teaching children. Like it holds in medicine that children are not diminished adults, and that paediatrics is not only curtailed internal medicine, it holds in educational process that adults are not augmented children in this process. It is true, that adult education methods are the same methods we could find in educational process of children, but the application of these methods is rather different (1). However, these methods are the same by the name but their forms, and educational tools and instruments used in the process, are different.

Basic terms in educational process

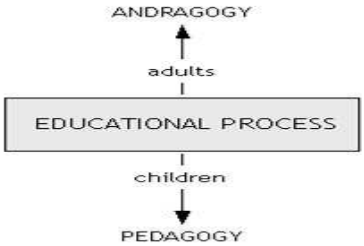
To properly understand the educational process, we need to distinguish between basic phenomena of this process, being education, teaching, learning, studying, and knowledge.

1. Education
   Education is described/defined as:
   - a process of teaching, training and learning, especially in schools or colleges, to improve knowledge and develop skills (2);
   - a long-term systematic process of gaining of knowledge, and developing of capacity, skills, habits, and personal characteristics (3,4).
2. Teaching
   Teaching is described/defined as:
   - giving lessons to students in a school, college, university, etc.; to help somebody learn something by giving information about it (2);
   - a process of introduction of students/learners to the process of studying, helping students/learners in this process, and reflective, methodological and methodical (systematic) guiding students/learners in their studying activity (3,4).
3. Learning
   Learning is described/defined as:
   - gaining knowledge or skill by studying, from experience, from being taught, etc. (2);
   - a direct changing/modifying of an individual with his/her own activity, that is being provoked by intrinsic (internal) needs, or extrinsic (external) incentives/stimuli/motives (3,4);
4. Studying
   Studying is described/defined as:
   - the activity of learning or gaining knowledge, either from books or by examining things in the world (2).
5. Knowledge
   Knowledge is described/defined as:
   - The information, understanding and skills that an individual gains through education or experience (2);
   - a system of captured facts, and generalization and development of abilities, skills and habits, that an individual captures more or less permanently, and puts to use (3,4).

Andragogy and pedagogy

The process of education of adults has its own scientific branch known under the term andragogy. On the contrary, the scientific branch that deals with education of children is known under the term pedagogy (Figure 1).

Figure 1. Scientific branches which support educational process of adults and children



1. Pedagogy

   The word comes from the Ancient Greek “paidagogeo”. This term is a complex term, composed of two basic terms “pais”, that means “child”, and “ago”, that means “to lead”. The term “paidagogeo” litterarly means “to lead the child". In Ancient Greece, “paidagogos” was a slave who supervised the education of his master’s children. This involved taking them to “didaskaleion”, that mean “school”, or a “gymnasion“, that mean “gym” (5). Thus, this term is related to educational process of children.

   Nowadays, there exist several very similar definitions of pedagogy, three of them being:

   - according to The Advanced Learner’s Oxford Dictionary (2), pedagogy is the study of teaching methods;
   - according to Merriam-Webster Online Dictionary (6), pedagogy is the art, science, or profession of teaching;
   - in TheFree Dictionary we could find among others that pedagogy are the principles, practice, or profession of teaching (7).

   These contemporary descriptions do not imply that pedagogy is related to children, but the name itself does.”

2. Andragogy

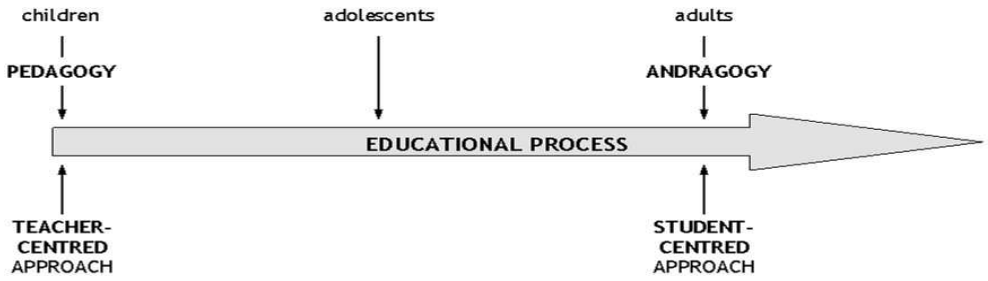
   Andragogy as a science has no long tradition, but andragogy as a process of education of adults is very old (8). The first use of the term “andragogy” - as far as it is known today - was found with the German high school teacher Alexander Kapp in 1833 (9). This term also, is based on greek terminology, despite this does not originate from Ancient Greece. It was constructed analogically as the term “pedagogy”. The first part of this term “andr-“, is basing on Greek term that means “adult male” (8,10). Thus the term means “to lead the adult male”. Because of “masculinisation”, still stirs up disapproval (8). Also the relation to the leisure time activities was disputable. The term “andragogy” has today wider meaning, like the term “gymnasium”, and it is not in any case related only to the leisure time activities.

   This discipline was in the past more developed in Europe, especially in France, Yugoslavia, and Holland (10). Several former Yugoslavia experts from this field are very well known, among them Serbian expert Dusan Savicevic, and Slovene expert Ana Krajnc (8). On the other hand, we need to mention also Malcolm Knowles. He brought in the seventieth the concept of science of andragogy to the United States of America (US), and developed this concept in his way which is well known (11). There exist also several critiques of this model of educational process (12), but this is beyond the scope of this module.

   There exist more or less similar definitions of andragogy (8). Krajnc, for example, defined andragogy as the science of education and teaching of adults, while Savicevic as relatively independent scientific discipline within more general science of education, that research education and learning of adults from all aspects (8).

   Originally, pedagogy and andragogy were seen as two different and separated processes, but current theory sees them on a continuum, with pedagogy on one end and andragogy on the other, rather than two separated processes (Figure 2) (13).

Figure 2. The continuum of educational process, and position of pedagogy and andragogy



We will discuss some of differences, which distinguish both concepts later on, but in this place would be worthwhile to emphasize maybe the most important distinction between them: andragogic approach is a student-centred approach, while pedagogic approach is teacher-centred (11,13,14) (Figure 2).

Didactics

A term, which is also used very frequently in the context of education, is “didactics”. The Croat expert of didactics, Vladimir Poljak, defines in his book didactics as “the branch of pedagogy, which is studying general rules of education” (3). In fact, it is studying the effective applicability of these rules, effective ways and forms of education (15). The word itself has Ancient Greek origin. The word “didaskein” means “to teach” (3).

In andragogy, didactics is defined in a wider way than in pedagogy. In andragogy it is not limited only on institutional or formal education (15).

Capacity building

If we are presenting these methods as important for capacity building, we need to clarify also two terms, being capacity and capacity building (16-19):

- capacity is defined as the ability of individuals, organizations or systems to perform appropriate functions effectively, efficiently and sustainably. Capacity has various dimensions:
  - it is not static but is part of a continuing dynamic process,
  - it does not exist on it own, but is linked with performance,
  - it is an instrument for an individual, team, organization or system to achieve objectives and
  - capacity always contributes to sustainability.
- capacity building or development is the process by which individuals, groups, organizations, institutions and societies increase their abilities to perform core functions, solve problems, define and achieve objectives, and understand and deal with their development needs in a broad context and in a sustainable manner.

These definitions, adopted by UNDP and UNESCO, emphasize the continuing process of strengthening of abilities to perform core functions, solve problems, define and achieve objectives, and understand and deal with development needs.

Why teaching adults is not the same as teaching children?

If the person is involved in education of adults, it is essential for him/her to be familiar with characteristics of child versus adult person in educational process. Children and adults in educational process differ in many respects. Some of them are summarized in Table 1 (13,20-26).

Table 1. Some differences between adults versus children in the educational process (13,20-26)

Feature	Children in educational process	Adults in educational process
Previous experience	Children have no life or other experience when entering educational process. They base their knowledge on the information they get in educational process from the teacher. Adolescents could enter the process with some limited amount of life or other experience.	Adults already have (various amount) of life and other experience when entering educational process. In this process, they add new knowledge to existing one. Additionally, they compare new knowledge to the old one, or they recall an old knowledge and put it into new frame.
The aim	Children have no aim and goals in the educational process. They even usually not think about this perspective of the educational process. The teacher points out the aim of learning.	Adults usually have clear aim and goals of the educational process. If not so, they determine them with the help of a teacher.
Source of motivation	Because children do not have aim and goals, they do not have intrinsic motivation in the educational process. The motivation is almost exclusively extrinsic (marks, fear, punishment, rewards, passed exam, scholarship, etc.)	Because adults have clear aim and goals, they usually have clear intrinsic motivation in the educational process. They also have different kinds of extrinsic motivation (e.g. requirement for competence or licensing, an expected promotion, adaptation to job changes, compliance with company directives, etc.).
Level of motivation	Because motivation is almost exclusively extrinsic, the level is usually lower than in adults.	Because motivation is extrinsic and intrinsic, the level is usually higher than in children. High level of motivation is in some cases conditioned by self-financing of education.
Level of autonomy and self-direction	Usually low.	Usually high.
Level of control over learning process	Usually low.	Usually high.

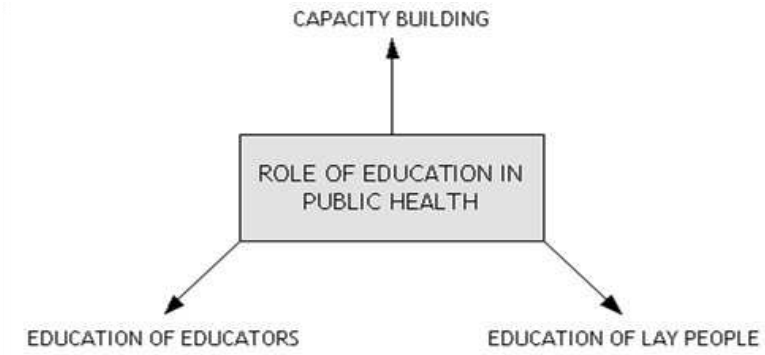
In short, the adult student in the educational process is usually characterized by following characteristics: he/she is a self-directed, self-motivated manager of personal learning who collaborates as an active participant in the learning process and who takes responsibility for learning (13).

Skills in principles and methods of educational process and public health

The role of educational process in public health

In public health, education has an important role. It takes place in capacity building as well as in public health profession itself in education of educators of lay people and in some cases the lay people as well (Figure 3).

Figure 3. The role of education in public health



For every profession, high quality and interesting educational process for capacity building is of utmost importance. This is more and more obvious also in public health (27). Public health workforce is becoming more and more important in establishing and/or maintaining public health care systems effective and efficient as much as possible (27). But is this workforce educated enough, especially if we consider the need for evidence-based public health, which needs highly educated and competent workforce in public health research? In US for example, only 44% of the public health workforce has formal public health education (27). European Commission Public Health Executive Agency (PHEA) states that many European Union (EU) member states and candidate countries have insufficient institutional and professional capacity for public health and that the process of reforming relevant services is slow. Compared to the US and other industrialized countries as well as some emerging economies, the relative lack of public health capacity in the EU is striking (28). But not only public health workforce is important. PHEA points out, that certain amount of public health skills is required for all kind of workforce involved in operating of health care systems (for example for medical doctors) (28). Skills should be of that amount that experts from professions involved in operating of health care systems could communicate between themselves in favour of end-users of these systems (for example patients).

Why in public health skills in andragogy are more important than skills in pedagogy?

In public health capacity building, which takes its place at universities, we meet different population groups, but majority of these population groups could be classified according to developmental psychology in early adulthood (29):

- adolescence: from puberty to about age 22-24 years;
- early adulthood: from about age 22-24 to about age 40-45 years;
- middle adulthood: from about age 40-45 years to about age 65 years;
- late adulthood: from about age 65 years to death.

If we take into consideration the developmental process of higher education in Europe (i.e. Bologna process), and Association of Schools of Public Health of the European Region, ASPHER, developmental recommendations for public health profession, we expect (or we already have) following target groups:

- students of the united first- and second-cycle (undergraduate and graduate) study of medicine and dental medicine. These students, which are mostly of age 19-25 years, are no longer children, but also not yet fully adults (23). In fact, they are finishing the adolescence phase, and are increasingly becoming young adults (23). The educational process of this population group is an exceptional art, demanding skills in pedagogy as well as in andragogy. It should pass from teacher-oriented to the student-oriented concept of education;
- students of the first-cycle (undergraduate, bachelor) study of public health or related studies, e.g. Health management study at Rijeka Medical School, Croatia (30), or Bielefeld University, Germany (31). This target population, which is of age 19-22, is very similar to the students of the first half of

study of medicine or dental medicine - they are finishing the adolescence phase and are becoming young adults;

- students of the second-cycle (graduate, master) study of public health (or branches of public health, e.g. health management, health promotion, etc.). These students are at least at the phase of younger adults (usually of age 22-24). Since the higher education in Europe is in the process of reorientation from full-time study to part-time study along to be employed, especially starting from master degree study according to Bologna process, older and older population is expected in this cycle. Working adults who will want to succeed in their profession will create a new majority among second-cycle students;
- students of the third-cycle (postgraduate, doctoral) study of public health or related studies are definitely adults of at least 24 years of age, and majority of them are employed;
- students of the professional third-cycle (e.g. postgraduate specializations of doctors in medicine or dental medicine) study of public health or related studies (of at least 25 years of age) are similar to the students of previous group. The principal difference between them is that this study is more professionally, and less scientifically oriented.

We meet almost only with adult population also when we use the educational skills in public health profession itself, irrespective they are used in education of educators, or in education of the lay people. Even if the end target group of public health messages is a group of children or adolescents, public health experts are not usually involved in educational process of these age groups by themselves. They mostly educate educators of these population groups.

This is the reason that we are emphasizing andragogical concept in this module, and not the pedagogical one. Adult students have unique needs, especially if they are employed. Before presenting these needs, we need to be acquainted with basic classifications and principles of didactical methods, irrespective they are used in andragogical or pedagogical process.

Teaching methods

Classification of teaching methods

Teaching or didactical methods could be classified in several ways, and these classifications could intertwine, since some methods could be seen from different perspectives (1,3,8,32-34). Ana Krajnc, the famous Slovene expert of andragogy, in her book presents three classifications: according to number of students, according to frontality, and according to formality of the presentation of teaching matter (Figure 4) (1).

Classification according to the number of students

The number of students certainly influences the choice of teaching methods used in educational process. It is clear that teacher would use a lecture rather than a consultation method when he/she has to teach a mass of 200 students at the same time. According to this criterion, methods are classified in:

- mass methods of education;
- group methods of education, and;
- individual methods of education.

In further considerations we will use this classification, thus it will be discussed in details later.

Mass, group, and individual methods can pass over one into another, depending on the educational objectives. When the end objective of educational process is adoption of general knowledge, or formation of general image/conception on the issue, then it is enough to use mass education methods. When the end objective is formation of personal viewpoints, and attitudes, group work is more appropriate. Practical exercising in small groups is especially recommended when end objective of educational process is adoption of practical skills (35). In all cases, educational process need to be supplemented by individual forms, at least guided individual study of provided recommended readings.

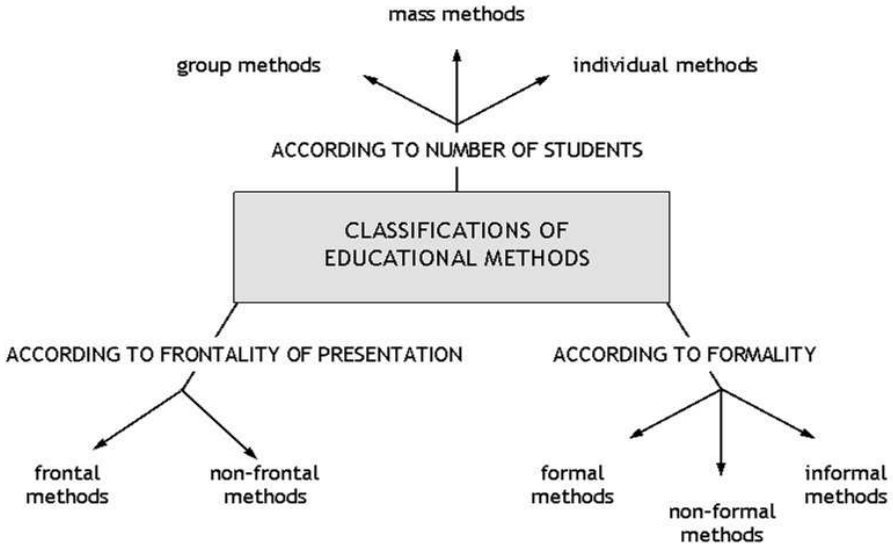
Classification according to frontality of the presentation of teaching matter

According to this criterion, methods are classified in:

- frontal methods of education – in frontal methods, teacher (a transmitter) transmits his/her knowledge to passive audience of students (receivers) (1). The term “frontal” is originating from the position of the teacher - he/she is standing in front of the audience while presenting teaching matter;

non-frontal methods of education - in non-frontal methods, students are active and they are at the same time receivers and transmitters of knowledge (1).

Figure 4. Three classifications of educational methods



Classification according to formality

Nevertheless the educational process is mostly understood as guided process of acquisition of knowledge, known also as formal, there exist also non-guided way. According to this criterion, methods are classified in:

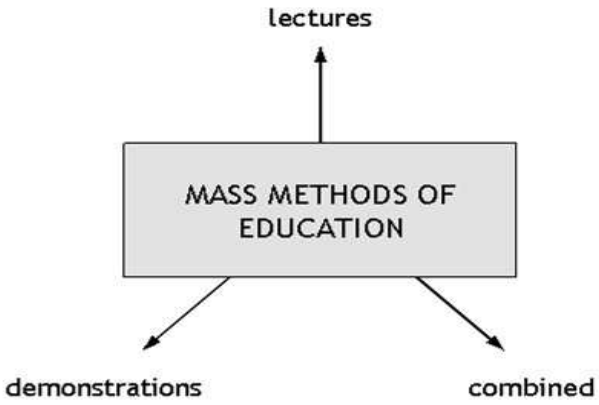
- formal methods of education – formal education refers to regular schooling that follows a normal pattern – admission students at about age six or over, promotion from grade to grade on a yearly basis, and use of a curriculum that covers a wide range of knowledge (19). It is directed towards acquiring of formally valuable and recognized qualification, level of education (from primary to the postgraduate), and title;
- non-formal methods of education – non-formal education refers to educational activities delivered to targeted groups, where there is a possibility to provide attention to individual learners. These activities may include for example courses and workshops that meet specific needs of society and it members (19). It is directed towards knowledge as educational achievement/result, and not as administrative one;
- informal methods of education – informal education refers to learning channels, such as mass media and mass publicity campaigns, where is little or no possibility for attention to the individual (19).

Mass methods of education

Mass methods of education could be first divided in formal and informal. Informal mass education is taking place all around us all the time through mass media, journals, books, internet etc. (1). This part of mass education is out of the scope of this module.

In formal mass education, usually frontal methods are used. The most well known method is lecture (1). Other well known and also frequently used method is demonstration (Figure 5).

Figure 5. Some mass methods of education



1. **Lectures**

According to Bergevin et al, the lecture is a well-prepared oral presentation of a subject by a qualified person (36). Lectures are basic teaching method in higher education (37).

This teaching technique, like others, has its strengths and limitations (36-38). Among strengths we could classify:

- can facilitate transmitting a message across to an audience of most any size, especially they are useful for mass education and education of large groups;
- saves time (particularly of teachers);
- saves money;
- gives feeling of safety to students;
- presents factual material in direct, logical manner;
- stimulates thinking to open discussion, etc.

Major limitations are common to all frontal methods of teaching:

- audience is passive (only reception);
- attention and acceptability of students varies;
- communication in usually one way, unless teacher stimulates students to ask questions;
- it does not permit assessment of progress of students;
- it does not permit individual tempo of learning;
- experts are not always good teachers, etc.

Lectures to be effective need clear introduction, content limit and summary.

According to Krajnc (1), lectures could take different forms/types, like:

- explication;
- narration, story telling;
- explanation;
- interpretation;
- description
- argumentation, etc.

Usually, two or more of them are combined at one lecture.

Krajnc, Pavlekovic and Levine give some ideas for better effectiveness of lectures (1,26):

- the lecture should be of maximum of 90 minutes of duration;
- the lecture should be organized, planned ahead;
- the content of the lecture should be logical in order of explanation;
- objective of the lecture should be clear;
- not too much information should be put in one lecture not to transgress the students' threshold of information uptake;
- periodic breaks that relax students by informally discussing the ideas that have been presented should be encountered;
- for keeping students' attention, examples, anecdotes, etc., should be included;
- visuals – charts, slides, and similar should be used, to allow students to see what they have been told;
- it is recommended to allow students for questions;
- it is also recommended to provide opportunities for small group discussion if possible, etc.

In adult learning, lectures could be only the introductory method, preparing students for more active work in small groups (discussion) or even individual (dialogue), which will be (some of them) discussed later on (1). Adults appreciate active involvement.

2. **Demonstrations**

By using this method, teacher shows how to perform a certain task (32).

Demonstrations are usually classified into two forms/types (26,36):

- method demonstration – illustrates how to do something in a step-by step procedure;
- result demonstration – shows the results of some activity, practice or procedure through evidence that can be seen, heard, or felt.

Group methods of education

Groups could be larger or smaller. Average large groups are as large as 20-25 students (1,35). This size of groups is regarded as being large and methods used are different than are used in small groups, which are on average large as 6-7 students (35). For using some teaching methods, even smaller groups are necessary. In larger groups different methods, usually of less active approach, are applied than in smaller groups (1,35).

Working with large groups

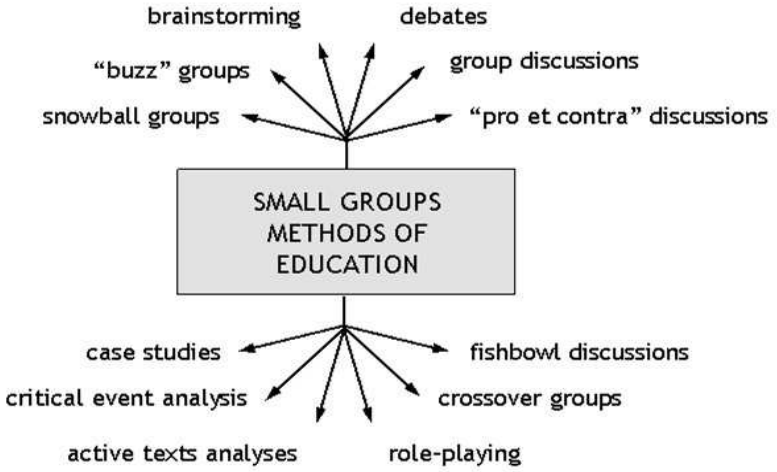
Methods used in large groups are very similar to that, used in mass adult education, being lectures and demonstrations (1). The only difference is that these methods, used in groups, even as large as 20-25 students, are more personal. Mass lectures, given for a mass of students as large as 150-200 students or even larger, are as effective as lectures given on television. The lecturer is far distant from the audience, almost anonymous. He/she has no personal contact with the audience, and they should be even supplied with loudspeakers (1).

Working with small groups

Teaching methods for small groups are much different than for large groups. They involve students actively. Teacher is in the role of coach and facilitator, giving students instructions how to work and what are the goals of their work (35).

Classical method used in education of medial doctors are laboratory and clinical exercises, but there exist also many other techniques (26,35,39), which could be very useful in public health education process. On this place only some of them could be presented (Figure 6).

Figure 6. Some small groups methods of education



1. **Text/document analysis (1)**
   A key text/document on topic under discussion is given to students to be studied. Each group could study the text from a specific point of view.
   This is excellent way to introduce a new topic or new concept for warming-up students before using other more pretentious methods of active learning.
2. **Group debate (1,32)**
   Group debate is a less pretentious active teaching method. In activity the teacher, as well as students are involved. In debate, each participant gives certain part of the topic under discussion as an answer, complement, or explanation to the previous speaker in the debate, or several of them. Speakers need continually pay attention to other speakers'. In mind, they analyze his/her/their speeches, and search for something not told yet, or not enough explained yet. At the end, the teacher supplements the missing part, if necessary.
   Despite this method is one of the oldest methods of education, it was neglected in pedagogical didactics, since for performing this method, students need to have certain amount of preliminary knowledge and experience, and enough of self-confidence to participate in the debate.
   In andragogy, this method could be very effective in strengthening and consolidating studying material, as well as for gaining new knowledge.
3. **Brainstorming (26,32,35,36)**
   Brainstorming is a method of free generating, outpouring and sharing of ideas related to the topic under discussion between students. The process has three phases.
   - in the first phase, students are encouraged to employ creative thinking and voicing their ideas. All ideas are accepted at the beginning of the process and no response, regardless of how useless or impractical, is omitted. Original ideas are very welcome;
   - in the second phase, the explanation, and categorization of ideas takes place;
   - in the third phase, the ideas are analyzed, evaluated, combined, refined etc.

This method can be excellent to help a group of students think creatively. It is very useful in situations when new ideas related to issue of interest are wanted.

Despite its usefulness this method has limitations. It is time-consuming and it is limited to the abilities of the participants. Also teacher's professional skills are important.

4. **Group discussion (1,26,35,36)**

According to Bergvin et al. (cited by Seaman) (36), group discussion is purposeful conversation and deliberation about a topic of mutual interest among 6-20 participants under the guidance of a trained participant called a "leader".

Discussion is not any more only presenting (like in lectures or in debates), or mutual supplementing of teaching matter. It is harmonization of contradictory standpoints, opinions, finding, statements, etc. of students about topic under discussion. Each student is confronted with the fact, that he/she must argue for his/her standpoint, opinion, etc., or justify it. It could also change it, if other students convince him/her opposite.

5. ***"Buzz" groups* (1,35,39)**

"Buzz groups" technique is used when very concrete task is to be solved in very limited time. Pairs of students (i.e. neighbours in a teaching room) are formed to discuss a question set by a lecturer. This method increases participation and engagement of students.

This method could be also used in groups of students to discuss/express difficulties they would have been unwilling to reveal to the whole large group.

It could be used in groups in which students know each other very well, but it also could be used with a goal that students become less inhibited. It is easy to implement in any size of student group and in most teaching rooms, even in lecture hall.

6. **Snowball groups (32,34,39)**

Snowball groups could be seen as an extension of "buzz" groups. The second name of this method is "pyramids".

This method involves progressive doubling of small groups. First, a pair of students is addressed to perform an activity. Pairs join up to form fours, then fours to form eights. These groups of eight students report back to the whole group (usually one representative of this group is a reporter) at the plenary session.

The disadvantage is that students could become bored with repeating the same task. For overcoming this problem, a sequence of increasingly more complex tasks could be applied. Another disadvantage is that it is a little time-consuming.

On the other hand advantage is that this method allows students to think for themselves or in small groups before open discussion. It also ensures good participation in plenary discussion.

7. **Fishbowls (32,34,39)**

Fishbowl method (or aquarium) involves one group observing another, active group. For the purpose of this method, two circles are made. The active group is sitting in the inner circle. It discusses an issue or topic while the outer circle group is observing/listening the discussion. The outer group looks for themes, patterns, soundness of arguments, etc. in the inner group discussion, analyze the inner group's functioning as a group, etc.

The disadvantage is that outer group students could become bored. For overcoming this problem, groups could switch places. Another variant is that participant from outer group that wants to make a contribution can switch the place with one participant from inner group when he/she is silent. This can happen many times during the discussion.

8. **Crossover groups (39)**

In crossover groups method, students are divided into subgroups that are subsequently split up to form new groups in such a way that maximal crossing over of information is assured. For the purpose of this method, students are labelled by two codes.

If we use for example two sets of codes, being A, B, C and 1, 2, 3, students are labelled as A1, A2, A3, B1, B2, B3, C1, C2, C3, etc. By asking all As, Bs, and Cs to work together, and then all 1s, 2s, and 3s, students could be mixed in such a way that each student meets at least one member from all other primary groups.

This method is excellent for mixing people and information. It is simple to organise. The plenary discussion could be omitted.

9. **Case study (26,36)**

By case study method, students analyse specific situation that they may be facing in the future. They search for solving the situation that has been presented.

One of the very important uses of this teaching method is in assessment how much students have learned and how comfortable they will be in using their knowledge in solving problems in the future.

10. **Role-playing (26,35,36)**

In role-playing, students are provided with a situation in which they are involved, and a role to play. They are asked to adopt a role, and play it. Afterwards they are required to debrief the experience, analysing the scenario, discuss their feelings, etc.

This method is extremely helpful in gaining skills in interacting with other people, since it is an effective way to get students familiar with realistic situations. On the other hand, the major disadvantage is that shy students could find this method stressful. For overcoming this problem, a role is given to a group of students who work out the details, but the role is actually played only by one of them.

Role-playing could be performed in small groups, as well as a frontal role-playing in front of the total group.

11. ***"Pro et contra" discussion*** (34,35)

"Pro et contra" discussion is a special type of group discussion. It could be also partially seen as role playing.

Two groups of students are given opposing positions to defend an opinion, standpoint, etc.

12. **Critical event analysis (35)**

This method is basing upon very concrete personal experience of students. Each student presents his/her concrete event, related to the topic under discussion. On the basis of personal story in the group discussion concrete solutions could be gained.

Problem-Based Learning (PBL)

PBL was originally developed at the Faculty of Health Sciences at McMaster University around 1965. PBL is an approach to learning and instruction in which students tackle problems in small groups under the supervision of a tutor. In most of the cases, a problem consists of a description of a set of phenomena or events than can be perceived in reality. These phenomena have to be analyzed or explained by the tutorial group in terms of underlying principles, mechanisms or processes. The tools used in order to do that are discussion of the problem and studying relevant resources. Typically, tutorial groups range from 5 to 10 students and one tutor, who may not only be skilled in group process, but expert in content, or may be primarily a facilitator (40,41).

The core concept in PBL is that the process of problem identification and problem solving can be a powerful way of learning because it is active, goal directed, and germane. Problem identification opens the door to a personal investment in problem-solving.

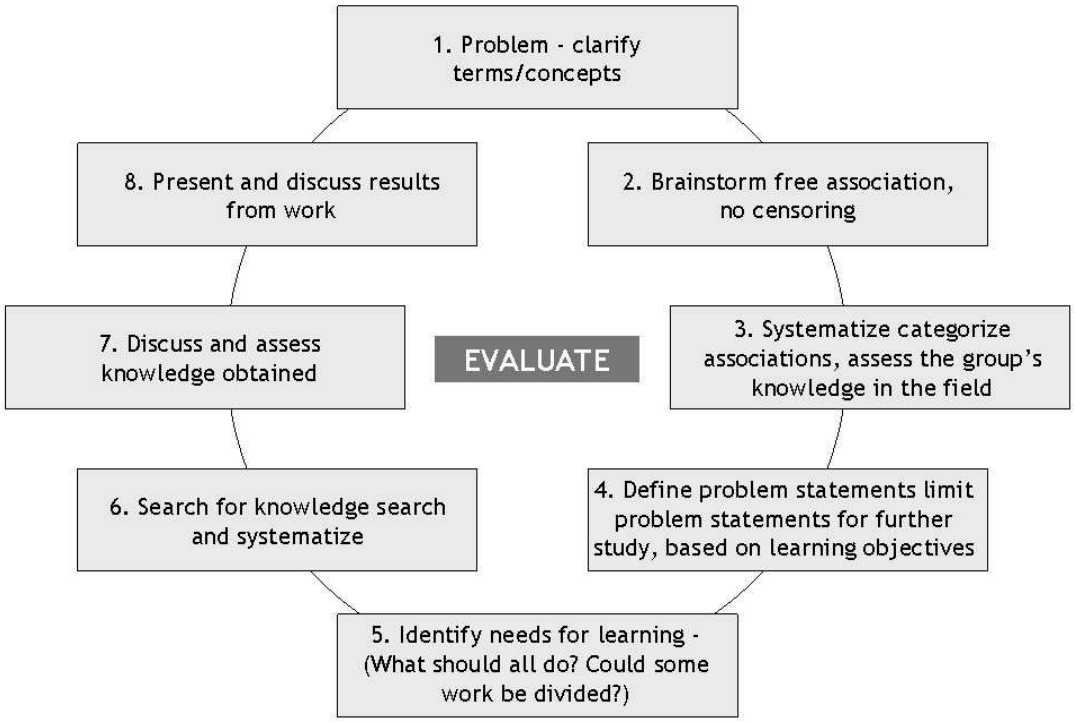
PBL has the following cognitive effects on student learning:

- activation of prior knowledge – the initial analysis of a problem stimulates the retrieval of knowledge acquired earlier;
- elaboration on prior knowledge through small group discussion, both before or after new knowledge has been acquired – active processing of new information;
- restructuring of knowledge in order to fit the problem presented - construction of an appropriate semantic network.
- learning in context. The problem serves as a scaffold for storing cues that may support retrieval of relevant knowledge when needed for similar problems.
- since students will tend to see problems presented as relevant and since they engage in an open-ended discussion, epistemic curiosity can be expected to emerge.

Even if there is some overlap, PBL differs from case-based learning. Illustrative cases are legitimate ways of enriching the learning process, but the method and goals differ from those of PBL. PBL cases require the steps of problem-identification, hypothesis formulation, analysis and validation.

PBL can be described as a cyclical process with several steps (Figure 7) (42). Group members must participate in all stages and movement between stages is not necessarily unidirectional.

Figure 7. The PBL circle. Adapted from Fosse, 2006 (42)



In Table 2, differences between traditional, teacher-directed learning and problem-based learning are presented.

Table 2. Differences between traditional, teacher-directed learning and problem-based learning

	For the student	For the teacher
Problems with traditional teacher-directed learning	Passivity Perceived lack of relevance of knowledge transmitted Often poor intellectual engagement Downloading of answers for which the students have not understood Little opportunity or incentive to conceive of questions that the teachers has not articulated Often lack of practice in kinds of collaboration which will be valuable in professional roles Relative lack of practice in setting personal learning goals, priorities, and methods	Tendency to use canned didactic talks and become jaded Frequently, lack of stimulation from fresh points of view and novel questions Relative lack of opportunity for cross-disciplinary faculty contact Often little opportunity for role modelling and mentoring, based on personal knowledge of students
Benefits with problem-based learning	Practice in actively identifying the main problems and issues in a case or other complex situation Practice in identifying and prioritizing helpful resources for learning Practice in systematic formulation of problem and hypothesis-setting Increased professionalism, separating vigorous intellectual discourse from personal relationships, as a member of a collaborative learning team Increased stimulus to learn underlying science and other principles as a result of relevant problem-solving	A satisfying sense of grater connection to student learning as a result of being in close touch with student thought processes A stimulus to rethink educational goals and methods Increased cross-disciplinary stimulation For many, a chance to learn or review up-to-date knowledge in fields close to one's own or otherwise interest

PBL places greater emphasis and responsibilities on students and tutors/teachers than the traditional teaching methods.

1. ***Students' responsibilities***

   In PBL, students' responsibilities are being present at the group meetings, being active at tutorials and between tutorials and between tutorials, takes responsibilities for moving through stages of the process, reflection upon and challenging principle ideas and concepts as they emerge in the group process.

2. ***Tutor's role***

   The tutor's role is one of moderator and facilitator. They are not there to lead the process or provide a definitive answer. They act as "sign post" and steer the process in a productive direction. Initially tutors contribute to identifying learning objectives and formulating problem statements, and ensure the chosen learning objectives are realistic. An effective PBL tutor will ask questions that stimulate students to work on a problem in greater depth. Further, the tutor should give advice on how to collect information. At last, but not least, contribute to solving group conflict.

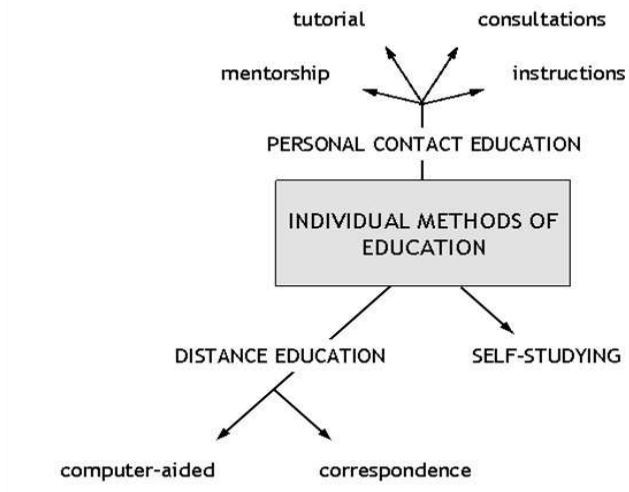
3. **Common responsibilities**

   Additionally, both students and tutor have the common responsibilities. They are responsible for the group functioning well, socially and academically. It is important that feed-back flow in both directions, both from the tutor to the students and from the students to the tutor. The group should agree at the outset some mutual expectations and ground rules. At the start of each subsequent session the group should reflect upon the group process.

Individual methods of education

By the term "individual methods of education", first association is on self-studying, but these methods are much more than only self-studying (1). In addition to self-studying, individual methods of education are also methods of distance education and methods of personal contact (Figure 8).

Figure 8. Some individual methods of education



Individual methods of education enable individualized process, oriented towards individual student (1,32).

Today, with explosion of communication technologies, distance education is very popular, and it is not any more limited on individual distance learning only. With new e-communication technologies, also group distance learning is possible. But this group of methods is out of the scope of this module. Because of its expansion it needs to be presented in a separate module.

Despite extensiveness of communication technologies, an important part of individual methods of education represents methods of personal contact education. In this type of education, teacher is in contact with one student at a time only (1). This part of the module will be dedicated to selected personal contact methods, being consultations, instructions, tutorials, and mentorship.

1. Consultations (1,26,36)

   Consultation is specific teaching method in a form of a mutual discussion, advising, clarification of certain problem, and answering questions. In consultation hours, students could also get information about recommended readings and similar. Consultations are not obligatory method - students seek contact with teacher voluntarily.

2. Instructions (1,26,36)

Instructions are very specific form of communication between the teacher, and the student, and it refers to narrower segment of educational process.

According to Dunn and Dunn (36), individualized instructions focus the instructional process on individual skills, abilities, interests, goals, and rate of learning. One of important features of this method is providing learning activities specifically formed to the needs, interests, and abilities of an individual student.

In comparison to consultations, instructions are one way communication. An instructor explicates the topic while student listens.

3. Tutorials (26,36)
   A tutorial educational method is performed when a single student needs specific help. The focus is usually the specific problems or concerns of the student.
4. Mentorship (1)
   In academic sphere, special method of education called "mentorship" or "mentoring" is applied for the most focused postgraduate study. In this educational relationship, a teacher is called "a mentor", and a student "a mentee". A mentor in this educational process is a trusted teacher who advise and guide his/her protege/protegee, how to learn and gain skills.
   Mentorship is to the certain extent similar to consultations, but between both methods there are huge differences. If consultation is a single educational episode, mentorship in opposite is a continuous process. It is a series of interrelated educational episodes, all directed to attain the educational objective.
   Mentorship is inevitably supplemented by intensive self-study of a mentee.

How effective are different teaching methods?

Before planning which didactical methods out of mass, group or individual methods of education would be used in educational process, and when, it would be useful to know how effective these methods are. This effectiveness is visualized by learning pyramid of the US National Training Laboratories, Bethel, Maine (43) (Figure 9). From this figure it is obvious that active methods of education, what mean the methods by which students learn with experience, have much higher retention rate than passive methods, especially those by which abstract learning is required. The far lowest retention rate is achieved at lectures. This traditional educational method is in Croatia and Slovenia still the basic method of education at academic level (37). The situation is very similar in other countries of South Eastern Europe. All this is true also for the public health education in these countries.

Figure 9. Learning pyramid with average learning retention rates (adapted from National Training Laboratories, Bethel, Maine) (43)

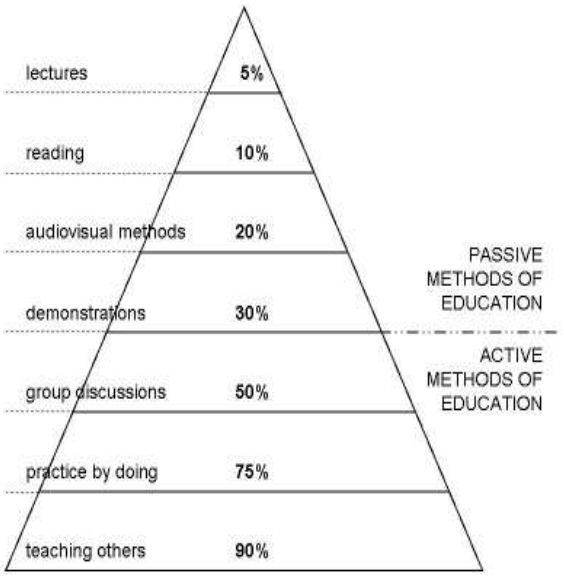


Table 3. Summary of basic educational methods

	Content	Characteristics	Carrying out
Lecture	Knowledge, important information, connection to the theory	Brief transfer of information, limited effects	At the first look easy to carry out
Exercise	Showing and practicing skills and procedure training	"Artificial circumstances", "out of common sense", often too simple	Organizational difficulties, equipment
Seminar	Description and analysis of case, setting down and solving the problems	Interesting, very often too much abstract, carry out as lectures	Well prepared students and very restrained teachers
Group work	Analysis of case, solving the problems, "difficult" situations	Learning from and support by peers, close interaction with teachers	Well prepared cases and problems as an impulse, clearly defined tasks, expensive
Practical work	Real professional life	Very useful learning, but not always systematic, can easily be superficial	Teachers as "role model" and leader, learning from experience
Reading	"Active" reading or content overview	Need for "critical" reading	First impulse and evaluation are important

In conclusion, it could be worthwhile to summarize basic educational methods. The summary is presented in Table 3.

Adult education process characteristics

Previously we summarized some differences between adults versus children in the educational process (Table 1), but also the process itself has different characteristics in andragogy versus pedagogy. At the beginning of this module we emphasized maybe the most important distinction between them: andragogic approach is a student-centred approach, while pedagogic approach is teacher-centred (11,13,14) (Figure 2). In andragogical process teacher is not a teacher in a classical sense anymore. In adult education, he/she is more a guide, a "lighthouse", a facilitator, a supervisor, than a "teacher" (1,13,22). Some other differences are presented in Table 4.

Table 4. Some differences of the process of adult education versus children education (22,23)

Feature	Educational process of children	Educational process of adults
The role of the teacher	The teacher's authority is very significant	Teacher is more a guide than a main source of information
Educational programme flexibility	Teaching programmes are based on educational programmes, which are based upon the educational requirements of the society. Teacher has little possibility to be flexible	Educational programmes are more flexible and oriented towards the ability of students, their previous experiences, their needs, etc..
Teaching methods	Mostly frontal methods of teaching are applied.	Active teaching methods are applied more frequently (group work, projects, discussions, etc.).

How to use educational knowledge in public health capacity building?

We are aware that for improving situation in the public health profession the shift from traditional passive educational methods to more active methods is strongly needed. Collins proposes in her paper ten tips in this process for radiologic education, but these tips could be used in public health as well (13):

- connect life experiences and prior learning to new information.
- involve participants in the learning process, serving as a facilitator and not just a supplier of facts.
- create educational programs that are organized with clearly defined elements, clearly showing how the program will help participants reach their goals.
- help learners see a reason for learning something by making it applicable to their work or other responsibilities of value to them.
- acknowledge the experiences that adult participants bring to the learning environment, allowing for opinions to be voiced freely.
- show learners how the learning will benefit them and create a comfortable and appropriately challenging learning environment.

- limit lecturing and provide opportunities for sharing of experiences, questions, and exercises that require participants to practice a skill or apply knowledge.
- accommodate different learning styles by offering a variety of training methods (e.g. group discussion, role-playing, lecturing, case studies, panel/guest expert, games, structured note-taking, individual coaching, demonstration, and variation in media used) and by using visual, auditory, and kinesthetic techniques.
- provide opportunity for feedback from self, peers, and instructor.
- promote group interaction.

We will illustrate now how to use this knowledge in practice using two case studies.

Case study

Case study 1: The course "Art of teaching in medicine: Croatian model of training in medical *and public health education"*

Needs for training of trainers: The story so far

In recent decades, we have been seen dramatic changes in the challenges facing the health profession. The nature and process of training for health workers has therefore undergone considerable change. Today, those involved in medical and public health education must adapt and find more innovative methods for teaching and learning what is now recognized as a much more complex subject embracing both the clinical and social paradigms associated with modern medicine.

Medical School University of Zagreb has recognized the need for changes in the way its medical curriculum is delivered and the need to improve the quality of medical study. Several activities have been implemented at Zagreb to meet those challenges:

1. a new curriculum has been developed and implemented based on the core competences for doctors working in the 21st Century identified from the academic and professional societies;
2. the introduction and development of innovative teaching and learning methods;
3. more opportunities for student feedback and evaluation of the curriculum;
4. formal training courses for those staff involved in delivering medical and public health education, particularly young teachers at the beginning of their professional careers, and quality assurance of the training provided.

Since 1990, the Department for Educational Technology at Andrija Stampar School of Public Health, has been responsible for the development and implementation of a programme of initiatives to improve the quality of medical education, Working on the project "The using video and computer technology in continuing medical education for Primary Health Care and Public Health", they had recognized the need for professionalism in medical education. A principle activity is to run a series of workshops, lasting three days, aimed at highly motivated and experienced teachers, to encourage the use of innovative concepts and principles in medical education.

As a result of this programme, the Croatian Association for Medical Education (CAME) was established in 1992, bringing together medical teachers from four existing Medical Schools in Croatia. The activities of CAME and the Department for Educational Technology at Andrija Stampar School of Public Health are mostly aimed at further training of medical and public health teachers. There are two main activities: formal courses covering basic of adult pedagogy and interactive workshops devoted to experienced teaching problems.

Formal course entitled "The Art of Teaching in Medicine" are aimed for younger teachers who are just entering the field of medical and public health education. They are organized twice a year for a group of about 25 participants and the educational methods are mostly learner-centred. The participants are provided by diploma and it is an important prerequisite for teacher's advancement.

The workshops which last 2.5 days are aimed to experienced teachers, they are subject focused and the main educational method is experiential learning. Following subjects have been discussed until now: curriculum development, a lecture - the oldest craft of teachers, problem-based learning, learning from experience, cognitive psychology and medical education, continuing professional development, student's - teacher's relationship, assessment in education, etc.

The both activities are very well evaluated, they are complementary, but the workshops are connected with higher motivation of participants and are considered to be more important strategy for professional development of teachers in biomedicine and public health.

***"The Art of Teaching in Medicine"*: Aims and objectives**

The course aims are to:

- encourage and enable medical and public health teachers early in their career to learn and put into practice contemporary approaches in teaching and learning medicine;
- understand the contextual factors influencing this process;

- encourage and motivate teachers to improve the quality of their work, and
- share experiences with others.

It gives a participant:

- a basic understanding of the concepts of study for health professionals, including challenges and dilemmas in teaching and learning;
- a theoretical framework for understanding factors influencing the quality of teaching-learning process;
- a range of evidence-based strategies, both traditional and innovative methods, to use in medical and public health education;
- a framework for planning, implementation and evaluation in the field of education;
- awareness of ethical issues relating to medical and public health education.

Participants: To whom the course is addressed?

Since 2000, the course was attended by more than 200 participants, mainly staff at Medical School Zagreb. All participants have been young or early career teachers from a range of different medical disciplines (clinical medicine, out-patient care, basic sciences, and public health). Since 2006, the course has been made compulsory requirement for staff involved in medical and public health education within the faculty, and part of the career development process for assistant professors.

Working format and main contents

The Course consists of six learning days with morning and afternoon sessions. A typical daily programme is presented in Table 5.

Table 5. A typical daily programme at the course entitled "The Art of Teaching in Medicine" at Andrija Stampar School of Public Health, Croatia

Time	Topic
A. 8-9.00am	Self-directed study (readings tasks related to the daily content)
B. 9-10.30am	Introductory topic, short lecture, discussion and exchange of practical experiences
C. 11-13.00pm	Exercise, small group work, tutorials, demonstration
D. 16-17.00pm	Panel discussion (invited guests – students, patients, medical school Board representatives-dean, vice-deans etc.)

In Box 1, there are listed all topics by day of the Course.

Box 1. Topics at the course entitled "The Art of Teaching in Medicine" at Andrija Stampar School of Public Health, Croatia, by day

Day 1:
- A. Introduction to the course. Expectations. Pre-test (MCQ)*
- B. Basic Problem-based learning, problem solving
- C. Educational strategies and effectiveness of medical and public health education
- D. Panel discussion: Europe and the World: Challenges in education for health professionals

Day 2:
- A. Self-study, preparation for daily topics
- B. Curriculum planning and modules development, developing learning objectives
- C. Teaching and learning tools: Module planning and implementation
- D. Panel discussion: Student as a partner in teaching-learning process

Day 3:
- A. Self-study, preparation for daily topics
- B. Basic educational tools in basic, clinical and public health sciences
- C. Bedside teaching/Teaching community health
- D. Panel discussion: Patient/community as a partner in teaching-learning process

Day 4:
- A. Self-study, preparation for daily topics
- B. Traditional vs. Innovative methods in teaching medicine: Lecture
- C. Lecture delivering (using video for self and peer feed-back)
- D. Evaluation

Day 5:
- A. Self-study, preparation for daily topic
- B. Teaching aids (video and computerized technologies, handouts, textbooks, simulators, skill laboratory, e-learning etc.)

C. Principles and methods in assessment
D. Panel: Assessment and exams as a part of teaching-learning process

Day 6:

A. Final test (the same MSQ)*
B. Staff development and academic standards
C. Panel: Strategies and actual policies of high education. Perspectives of Croatian Medical Schools development.
D. Course evaluation and future plans

Process and impact evaluation

Process evaluation

On the last day of the programme participants are asked to evaluate the course, using quantitative and qualitative methods (written questionnaire and group discussion). Summary of findings are as follows:

- according to the feedback the course fulfilled the expectations of participants (mean grade 4.0±1.2, ranking from 1 the lowest to 5 the maximum grade),
- all participants confirm it has been useful and has benefited their everyday teaching process (mean grade 4.2±0.9),
- those elements of the course scoring most highly in the evaluation were:
  - builds on and expands their previous teaching experiences,
  - directing interests to teaching,
  - stimulation of exchange of experience with colleagues from different departments,
  - peer discussion as a continuous support for their advancement in teaching,
  - satisfaction in communication and sharing doubts with experienced teachers,
  - panel discussions with guests and an overview of present situation in medical education.
- those aspects of the course evaluated less highly and where improvements could be made:
  - insufficient participation of "older teachers" and key decision makers within the faculty,
  - doubts as to whether implementation of skills mastered in the course will transfer to practice.

The course impact

The course impact was analyzed using students' assessment of all medical teachers within the Medical School Zagreb. Student feedback was compared with their mean grade, with those who passed the course, and their views as to how many of the teachers implemented innovative teaching and learning (modules) within the medical curriculum. Summary of findings are as follows:

- the mean grade for all faculty (overall assessment by students on scale 1 the lowest to 7 the maximum grade) is 5.1±2.3, for young or early career who had participated in the training programme, this was higher at 5.9±1.2,
- during the last seven years 15% of the trained faculty has introduced innovative modules in their teaching practice, mainly in electives.

Lessons learned

In comparison with other training for trainer's courses in medical education, with similar aims and objectives, this programme has at least two positive characteristics or strengths:

1. Trainers are not experts in pedagogical or didactical knowledge and skills. They are well experienced medical doctors/teachers aware to share their own experiences and reflections with participants.
2. Putting together a mixed group of participants (with different pre-knowledge, field of work, interest's expectations and teaching environment) has a positive influence on their motivation and encouragement.

Beside positives, at least two dilemmas remain:

1. Do we need (formally) certificated medical teachers or voluntary contribution of those who love "the art of teaching medicine", and will personally and with enthusiasm invest in their own development?
2. Impact evaluation (positive assessment of those who finished the course) although positive using students' questionnaires can be biased and their response could be influenced by factors not directly related to the training programme (for example student expectations, teachers personality, student learning styles etc.). Further evaluation and research is needed to expand and clarify the findings from this study and influence the development and implementation of training programmes in medical education.

In summary, the CAME training course in innovative teaching and learning is shown to be important for creating a new culture of teaching and learning in medicine and public health. It has been particularly successful in increasing awareness and motivation of the young or early career members of medical faculty for mastering the fundamental skills in medical education. However, participation of older, more senior or

long-serving teachers are somewhat limited, resulting in the current variation in the learning experiences of medical students. The ongoing availability of CPD workshops in innovative teaching and learning in medical education is essential for both the development of specific skills for application of acquired principles in everyday practice as well maintaining the motivation and enthusiasm of those responsible for medical and public health education.

Case study 2: Teaching statistical methods in public health on postgraduate public health course Ljubljana medical faculty, Slovenia

Introduction

Usually, all users of statistical methods with less mathematical background, especially those who use them only occasionally, meet big problems when using these methods. Since these methods represent very important tool in research in public health, it is very important to enable public health students the learning approach based on comprehension of statistical methods not only using them as a "cook book".

For this aim to be attained, learning must be approached with consideration. It is undoubtedly important:

- the teaching matter to be distributed logically and units should be smoothly associated one to another;
- not to use too many equations or mathematical expressions, especially not in professions with less abstract way of thinking (in medicine in particular);
- to enable the students to use the methods in practice, meaning transferring the emphasis from the lectures to practical work on an example (learning with experience);
- to watch carefully all the time if the students are following the explanation and practical use of the teaching matter; the teacher must strive for not concluding a subject until it is accepted by the majority of students;
- to use new pedagogical approaches at lecturing, for example working in smaller groups, learning with discussion etc., whenever possible;
- to use adequate tools, meaning learning with the statistical program of good quality; nevertheless certain calculations shall be carried out manually or partly manually, particularly where this shall be necessary for the procedure to be easier understandable;
- students shall be proposed, if not supplied with, corresponding studying material.

The main purpose of teaching public health students statistical methods with comprehension is for them to be able to have entire process of a public health research, from the definition of a problem as a starting point, through collecting the relevant data and analyzing them, finally to interpretation of the results of the analysis under control. This does not necessarily mean that they will need to do the entire process by themselves. Lack of general acquaintance with research tools including statistical methods most often results in a public health worker, not be able to translate the results of an analysis into technical language of his or her profession. This is usually followed by a very complicated interpretation, which can be very hard to understand, especially to decision makers and politicians to whom the results are intended for. In other words, teaching statistical methods with comprehension means to make future public health experts to be able to speak the same language as statistic experts, who are responsible for technical realisation of the methods.

Process of teaching

In teaching statistical methods in public health at postgraduate level of education in Slovenia we try to follow the principles just described as much as possible. It seems that this could be the right way to achieve the aim already described.

In continuation, a short description of the process of teaching statistical methods at the 2 semester (60 ECTS) postgraduate course of public health, which is a compulsory part of the professional third-cycle study of public health postgraduate specialization in public health for doctors of medicine or dental medicine (for others is optional), is given. This method is used now successfully for about 10 years.

Educational objective

End objective of educational process is adoption of practical skills. For achieving this objective, practical exercising in small groups is especially recommended (35).

Target group

Target group are adult postgraduate students of 25-40 years of age, occasionally some students up to 55 years of age are expected. According to developmental psychology classification, this target group is classified in early to middle adulthood (29), thus andragogical principles are to be used.

Educational programme

The experiences of long standing show that the clear elements of teaching matter in statistics, and

right sequence of these elements are crucial for success.

We conceived the whole teaching process on the relationship analysis, which is situated centrally. Accordingly, teaching starts with theoretical concept of relationship analysis. This is followed by construction of thinking process from the most simple building blocs of statistics – basic concepts and methods, gradually to more complex ones. When the technique is absorbed, we return to the relationship analysis. The realization of lessons is designed in modular form:

1. Introduction:
    - MODULE 1: Introduction to relationship analysis in public health, and use of statistical methods (objective: understanding basic concept of relationship analysis);
2. Principles of statistical methods:
    - MODULE 2: Statistical concepts and statistical describing of data (objective: understanding of basic statistical concepts, sample distributions and their meaning in statistical inference);
    - MODULE 3: Statistical inference (objective: getting familiar with techniques of statistical inference using worked examples);
3. Using statistical methods in relationship analysis:
    - MODULE 4: Principles of relationship analysis – univariate analysis (objective: transfer of basic principles of statistical methods to simple relationship analysis; students understand, and are able to perform analyses independently);
    - MODULE 5: Principles of relationship analysis - multivariate analysis (objective: extending the methods for simple relationship analysis to more complex ones, getting familiar with principles of multivariate relationship analysis; understanding technique of frequently used statistical methods in public health, i.e. logistic regression; students are able to perform analyses under the guidance of the teacher);
    - MODULE 6: Specific designs (e.g. repeated measurements analysis).

Based on our experiences, students must be given enough time for basic concepts. It is hard to say how much time would this take since groups of public health students are differing considerably. At the beginning of single course, the groups of students mostly are possible to be of very heterogeneous background knowledge. A good teacher shall be able to carry out the whole teaching matter with no regard to the emphasis laid to individual methods.

Teaching process

Teaching method

For successful outcome of teaching process it is essential to use teaching methods with emphasis on the active involvement of students in the process. The approach of learning with experience in small groups based on global worked example using high quality computer programme is used, combined with large group methods, being introductory lecture and demonstration (method demonstration, as well as result demonstration). Occasionally, “buzz” groups method is used.

Size of groups

In a large group, there are 12-18 students.

In each small group there are no more than three students with optimal number of two. Every single small group of students, together with a computer constitutes one working unit. The students are distributed considering following criteria, if possible:

- one student in a group is familiar with working on a computer;
- one student is not a beginner of a postgraduate specialistic study of public health.

This approach enables students to actively share their knowledge and help themselves inside each small group.

Other active involvement

Additionally, discussion inside groups is stimulated, as well as discussion between groups. Also, the comparison of results between working units, whenever possible, is stimulated.

Data material and teaching tools

Data material

For already several years at teaching statistical methods in public health in Slovenia data collection which enables learning such methods in quite a pleasant way has been used. These are the data collecting within the Perinatal Informational System of Slovenia (PISS) (44), which is considered to be one of the permanent collections of medical data of the highest quality with the many years' tradition. It was started in 1987 when collection of data started on a uniform form in all Slovene maternity hospitals.

Data material for teaching is only a small piece out of the whole collection PISS, prepared especially for this purpose. Safeguard of personal data is assured so that all personal identificators have been removed. Moreover, only the data are selected from the whole collection which shall be used for the teaching purpose.

Special attention was paid to include such data, which are possible to play the role of the "effect" as well as the data which can play the role of "cause". Also, they are chosen to have diverse statistical characteristics as to be able to represent the majority of statistical methods.

The basic data material for all statistical activities is composed of more than 6000 statistical units, representing the model of a population. For teaching different statistical methods, samples of various sizes are then randomly selected from the population database. The smallest sample is composed of 30 units and the largest one of 800 units. The students shall select some of such samples from the population while performing practical work by themselves and others shall be prepared by the teacher. First ones are, on principle, different at each working place and shall be, at the end of the corresponding exercise, deleted by the students or by the teacher. Others are the same for all working places, constantly present there and are used for many times. Data material is composed of the following permanent databases:

1. Population database;
2. Database with a sample of 100 units – main database for teaching principles of statistical methods;
3. Database with a sample of 500 units – additional database for teaching principles of statistical methods;
4. Database with a sample of 800 units – main demonstrational database and main database for teaching multivariate statistical methods.

Notwithstanding the number of units in a single database, all databases are including the same statistical variables. On this data, most of statistical methods could be presented in an obvious way as well as numerous possible traps of their misuse.

Teaching tools

Adequate equipment is also very important for teaching statistical methods in a proper manner, which is carried out mostly by the means of a computer with a statistical program of suitable quality.

At the University of Ljubljana, the statistical program SPSS (Statistical Package for Social Sciences) has been used for this purpose for already a few decades. At the beginning, it was possible to be used only on main university computer. However, for the last about ten years the University has the license to enable the distribution of the program within its members under favourable conditions for each personal computer. Two main purposes are so achieved: a program of good quality is used within the University, enabling the results to be compared within the university and as this program is widely spread also enabling the results to be compared to other countries.

However, whenever it is necessary for easier understanding of methods, the program can be supplemented by using a calculator and statistical tables.

Teaching material

In general, materials for modules are composed of:

1. Short summary of the subject covered by the unit
2. Exercises for practical work.
3. List of data files for practice.
4. Teaching material necessary for the tasks which are foreseen to be calculated manually.
5. Copies of the results of analysis carried out by the means of the SPSS program on a sample of 800 units.

Studying material

Studying material prepared for students, is composed of:

1. A short manual for working with the SPSS program, so that they were able to start working by themselves with a program from the very beginning.
2. Short summary of the subject, which was not intended as being the only source of statistical theory for the students, but to help directing their flow of thinking and as a help with practical work.
3. Collection of exercises for practical work.
4. Copies of the main results of the SPSS programme, so that the students are able to mark down their comments and explanations of printouts for helping them at studying.
5. Advised textbook for deepening their theoretical knowledge, which could be find in a Central Library of Faculty of Medicine or bought for a reasonable price.

Comments

Using described methods, we consider several principles of andragogy:

- students are actively involved in the educational process,
- teacher is a facilitator and not just a supplier of facts;
- use of life experiences and prior education is stimulated, as well as group interaction;
- feedback from teacher is given in few days.

The effectiveness of educational process is being evaluated through statistical analysis project, which is presented to other students.

Exercises

Introduction to Task 1: Micro-teaching exercise

Rationale

Lecture, the most often used method in teaching process, is only one of several ways of passing on factual information, giving an overview and motivating participants in teaching and learning process. Lectures should be carefully prepared, clearly structured and have defined SMART (specific, measurable, achievable, realistic and timing) aim and objectives. To lecture effectively requires that attention be paid to presentation skills and the use of any visual or technical resources.

Learning objectives

Learning objectives are:

1. Practice lecturing an 8-10 minute segment of a planned teaching exercise;
2. Analyze one's own and others' styles in a micro-teaching format.

Task 1

Participants:

- prepare an 8-10 minute lesson on topic of their own choice,
- they present the lecture to their colleagues using any format that they choose[27.] The lecture is videotaped. Both performance/presentation (verbal and non-verbal) is analyzed as well as components of the lecture.

It is important to have a constructive feedback from the audience. Colleagues may be able to give an objective critique and provide positive advice on content and delivery. On the other hand, personal reflection may be undertaken by personal viewing a video of own presentation.

Introduction to Tasks 2-4: Working with small groups

Tasks 2-4 are basing on following assumptions:

- we have a large group of 24 students. We split this large group to 4 small groups of 6 students,
- the main teaching matter is health promotion. Accordingly, all tasks are basing on teaching modules, published in the manual for teachers in PH-SEE network "Health promotion and disease prevention. A handbook for teachers, researches, health professionals and decision makers" (45),
- students have already read the theoretical background of this module.

Task 2

Each group is given a theme to be virtually taught from above mentioned manual (Table 6):

Table 6. Themes to be virtually taught

Group	Theme	Chapter in the manual
Group 1	Concepts and Principles in Health Promotion	1.1 Concepts and Principles in Health Promotion
Group 2	Healthy Public Policy	1.2 Healthy Public Policy
Group 3	Media and Health	4.3.2 Media and Health
Group 4	Smoking cessation campaigns	5.6.1 Quit and Win Campaign

Students are obliged to read the Theoretical background parts of proposed modules to be familiar with the contents. They can also read the Case study part, but they are supposed to omit the Exercise part.

[27] Important notes in lecture delivery:

1. Lecture is often arranged in three sections: introduction, body and conclusion.
2. Introduction provides motivation and enthusiasm to inspire the audience to concentrate for the main body of the lecture. It could be useful to provide a statement of the objective(s) of the lecture (e.g. „At the end of this lecture you should be able to....") and indicate the key points which are to be noted on the way (outline).
3. It is crucial to the success of the lecture that you maintain the audience's attention. Asking yourself the following questions will make this easier to accomplish:
   - Where should I stand?
   - How should I speak?
   - What should I do?
   - When should I change style?
   - How to communicate with the audience?
4. Any additional materials (illustrations, handouts, PowerPoint or other multimedia presentation) presented in a lecture must be of good quality.
5. Interruptions and technical faults provide a test of the lecturer's skills. It may be wise to work out in advance a strategy for dealing with some of the following before being faced with them: uninvited questions, participants falling asleep, technical failures, etc.

Task 3

Every group is proposed to use different teaching method for small groups to perform. In Table 7, the teaching methods to be used are listed.

Table 7. Teaching methods to be used

Group	Teaching method	Tips
Group 1	Snowball groups	Snowballing is started in this case from single student, than form twos, and at the end make plenary session.
Group 2	Fishbowl	Three students are in the inner circle, and three in the outer.
Group 3	Crossover groups	Students are labeled as A1, A2, B1, B2, C1, and C2. First all As, Bs, and Cs work together, and then all 1s, 2s, and 3s work together
Group 4	"Pro et contra" discussion	From the group of 6 students, two groups of 3 students are formed. One group prepares arguments in favour of public health campaigns, in the other group are the opponents.

Students at the plenary session perform indicated teaching methods in front other 18 students.

Task 4

From the manual (45), each group chose another module and propose a teaching method that is according to their opinion the most suitable. Afterwards they present to other students in short the theme they have chosen and the proposed teaching method. Other students discuss.

References

1. Krajnc A. Methods, forms, and means in adult education. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.101-221.
2. Oxford Advanced Learners Dictionary. Seventh Edition. Oxford: University Press; 2005.
3. Poljak V. Didactics (in Slovene). Ljubljana: Drzavna zalozba Slovenije; 1974.
4. Tomic A. Selected didactical topics (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani; 2001.
5. The Online Etymology Dictionary. Available from URL: http://www.etymonline.com/index.php. Accessed June 5, 2013.
6. Merriam-Webster Online Dictionary. Available from URL: http://www.merriam-webster.com/dictionary/medicine. Accessed June 5, 2013.
7. TheFreeDictionary. Available from URL: http://www.thefreedictionary.com/. Accessed June 5, 2013.
8. Licen N. Introduction to adult education (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za pedagogiko in andragogiko; 2006.
9. Reischman J. Andragogy. History, Meaning, Context, Function. Available from URL: http://www.uni-bamberg.de/fileadmin/andragogik/08/andragogik/andragogy/index.htm. Accessed June 5, 2013.
10. Smith MK. "Andragogy", the encyclopaedia of informal education. Last update 1999. Available from URL: http://www.infed.org/lifelonglearning/b-andra.htm. Accessed June 5, 2013.
11. Knowles MS, Holton EF, Swanson RA. The adult learner: the definitive classic in adult education and human resource development. 6th ed. Amsterdam: Elsevier, 2005.
12. Burge L. Beyond andragogy: Some explorations for distance learning design. J Dist Educat, 1988;3:5-23.
13. Collins J. Education techniques for lifelong learning. Principles of adult learning. RadioGraphics 2004;24:1483–1489.
14. Imel S. Teaching adults: is it different? ERIC Digest No. 82. Available from URL: http://www.ericdigests.org/pre-9211/teaching.htm. Accessed: June 5, 2013.
15. Krajnc A. Andragogical didactics. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.13-38.
16. UNDP. Capacity development resources book. New York: UNDP, 1997.
17. UNDP. Capacity assessment and development, in a systems and strategic management context. Technical advisory paper 3. New York: UNDP; 1998.

18. UNESCO. UNESCO thesaurus. Available from URL: http://databases.unesco.org/thesaurus/. Accessed June 5, 2013.
19. International Institute for Educational Planning. Guidebook for planning education in emergencies and reconstruction. Paris: UNESCO; 2006.
20. Krajnc A. Experience of adults and education. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost, 1979; p.69-82.
21. Krajnc A. Motivation of adults for education. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.83-100.
22. Ruzgiene L. Strategies and reflections upon adult education. In: Wie AL (editor). Improving adult learning. Nesna: Hogskolen I Nesna; 2003. p.14-22.
23. Jaksic Z. Student-centred teaching. In: Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The art of teaching in medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005. p.87-95.
24. Lieb S. Principles of adult learning. Available from URL: http://honolulu.hawaii.edu/intranet/committees/FacDevCom/guidebk/teachtip/adults-2.htm. Accessed: June 5, 2013.
25. Egle C. A guide to facilitating adult learning. Curtin ACT: Rural Health Education Foundation, 2007.
26. Levine SJ. Effective teaching. A set of papers focusing on the teaching of technical information to adults. In: Michigan State University. Approved trainer of registered technician self study manual. Available from URL: www.ipm.msu.edu/uploads/files/SelfStudyManual.pdf. Accessed June 5, 2013.
27. Bryan RL, Kreuter MW, Brownson RC. Integrating Adult Learning Principles Into Training for Public Health Practice. Health Promot Pract Health Promot Pract. 2009;10:557-63.
28. European Commission Public Health Executive Agency. Conference on professionalization and capacity building in public health in South-Eastern and Eastern Europe: the legal and educational framework, Luxembourg, 20/21 February 2008. Draft report. Luxembourg: PHEA, 2008.
29. Zupancic M. The subject and history of developmental psychology. In: Marjanovic-Umek L, Zupancic M, editors. Developmental psychology (in Slovene). Ljubljana: Znanstvenoraziskovalni institut Filozofska fakultete; 2004. p.6-27.
30. School of Medicine Rijeka. Study of organization, planning, and management in health care (in Croatian). Available from URL: http://www.medri.hr/studiji/organizacija/o%20studiju.htm. Accessed June 5, 2013.
31. University of Bielefeld, Faculty of Public Health. Bachelor of Science in Health Communication. Available from URL: http://www.uni-bielefeld.de/gesundhw/english/academic/bhc/index.html. Accessed June 5, 2013.
32. Loginova N. Adult education methods and methodologies. In: Wie AL, editor. Improving adult learning. Nesna: Hogskolen i Nesna, 2003. p.30-42.
33. Piderit GA, Quijano LF. What are the main techniques used in adult education? Available from URL: http://fcis.oise.utoronto.ca/~daniel_schugurensky/faqs/qa14.html. Accessed: June 5, 2013.
34. Puklek Levpuscek M, Marentic Pozarnik B. Work in small groups for active study (in Slovene) Ljubljana: Filozofska fakulteta, Center za pedagosko izobrazevanje; 2005.
35. Vrcic-Keglevic M. Learning in small groups, seminars. In: Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005. p.87-95.
36. Seaman DF. Adult education teaching techniques. College Station, TX: A&M University; 1977.
37. Pavlekovic G. Lecture. In: Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The Art of Teaching in Medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005. p.145-151.
38. Cantillon P. ABC of learning and teaching in medicine: Teaching large groups. Br Med J. 2003;326:437-440.
39. Jaques D. ABC of learning and teaching in medicine: Teaching small groups. Br Med J. 2003;326:492-494.
40. Borrows HS, Tamblyn RM. Problem-Based Learning. New York, Springer Publishing Company; 1980.

41. Schmidt HG. Foundations of problem-based learning: some explanatory notes. Medical Education. 1993;27:422-32.
42. Fosse M. PBL. In: ETC-PHHP/EUMPH Summer School: Strategies for Health in Europe. Readings. Galway: ETC-PHHP; 2006.
43. Hall JJS B. Teaching Methods and Retention [Homepage on the Internet]. Available from URL: http://www.simulations.co.uk/pyramid.htm. Accessed June 5, 2013.
44. Institute of Public Health of the Republic of Slovenia. Perinatal Informational system of the Republic of Slovenia. Available from URL: http://www.ivz.si/Mp.aspx?ni=46&pi=5&_5_id=56&_5_PageIndex=1&_5_groupId=185&_5_newsCategory=&_5_action=ShowNewsFull&pl=46-5.0. Accessed June 5, 2013.
45. Donev D, Pavlekovic G, Zaletel-Kragelj L, editors. Lage: Hans Jacobs Publishing Company, 2007. p.378-395. Available from URL: http://www.snz.hr/ph-see/Documents/Publications/FPH-SEE_Book_on_HP&DP.pdf. Accessed June 5, 2013.

Recommended readings

1. Cantillon P. ABC of learning and teaching in medicine: Teaching large groups. Br Med J. 2003;326:437-440.
2. Dent JA, Harden RM, editors. A practical guide for medical teachers. London: Elsevier; 2005.
3. Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The art of teaching in medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005.
4. Jaques D. ABC of learning and teaching in medicine: Teaching small groups. Br Med J. 2003;326:492-494.
5. Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979.
6. Wie AL (Ed.). Improving adult learning. Nesna: Hogskolen i Nesna, 2003.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Designing and planning educational programmes in public health**
Module: 2.40	ECTS (suggested): 0.3
Author(s), degrees, institution(s)	Gordana Pavlekovic, MD, PhD, Associate Professor Andrija Stampar School of Public Health, Medical School, University of Zagreb, Croatia Lijana Zaletel-Kragelj, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Anja Kragelj, Student of Andragogy at Department of Pedagogy and Andragogy, Faculty of Arts, University of Ljubljana, Slovenia Natasa Skerget, Student of Pedagogy at Department of Pedagogy and Andragogy, Faculty of Arts, University of Ljubljana, Slovenia
Address for correspondence	Gordana Pavlekovic Andrija Stampar School of Public Health Rockefellerova 4, 10000 Zagreb, Croatia E-mail: gpavleko@snz.hr
Keywords	Adult education, educational programme, planning.
Learning objectives	After completing this module students should: • know the definition and characteristics of educational programme; • be familiar with different models of educational programme planning; • be familiar with educational programme planning process.
Abstract	Programme planning is a delicate phase in the educational process. If by education we mean the provision of systematic organized learning experience, programme planning is the core phase. The module is describing principles of educational programme planning and provides some skills in the planning process through presenting case studies.
Teaching methods	An introductory lecture gives the students first insight into the educational programme process and its characteristics. After introductory lectures, students first carefully read the recommended readings. Afterwards they form small groups. Every group is given a problem concerning planning of some public health issue to be taught. In each group, students try to plan a programme and present it to other students. Other students critically discuss their plan.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room, rooms for small-group work; • equipment: computers (one computer for 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Elaborated example of a plan for educational programme.

DESIGNING AND PLANNING EDUCATIONAL PROGRAMMES IN PUBLIC HEALTH

Gordana Pavlekovic, Lijana Zaletel-Kragelj, Natasa Skerget, Anja Kragelj

Theoretical background

Basic terminology

Before discussing the process of planning educational programmes, it would be worthy to clarify terminology in this field. Two terms are used in relation to planning an educational process, being "educational programme" (or "program") and "curriculum" (1,2). To understand them, we need first to clarify some other terms.

Programme

Several definitions exist on what the term programme means, among which we can find the following:

- a formal set of procedures to conduct an activity (3);
- a set of projects designed to achieve common, long-term goals (4);
- a set of organized but often varied activities (a programme may encompass several different projects, measures and processes) directed towards the achievement of specific objectives (5)
- a generic term taking into account a collection of coordinated projects/sub-programmes, single measures, processes and/or services aimed at achieving a common aim. The key elements of a programme include its inputs, processes, outputs, outcomes, impacts. It is defined and limited in terms of time, scope and budget (6).

In general, a programme is a planned sequence of activities designed to achieve specified objective. It has several steps, being needs assessment, programme planning, programme implementation, and programme evaluation:

- needs assessment - this is a step at which information about community health problems are gathered. On the basis of these information assumptions on how the needs/problems could be addressed, and the objectives/goals that should be reached are set up;
- programme planning - planning step - uses the assumptions set up at the previous step to plan a programme of activities;
- programme - implementation step - refers to the follow-up of the activities in accordance with the plan. Implementation could be expressed in terms of operational or action plans which commonly outline concrete activities, time frames, responsibilities, budgets etc., for the achievement of different objectives of the programme;
- programme evaluation - when running a programme, we want to know how far the programme went and how effective it is in achieving its goals/objectives. We are able to answer these questions by performing the so-called programme evaluation process. A programme evaluation is a periodic review and assessment of a programme to determine, in light of current circumstances, the adequacy of its objectives and its design, as well as its intended and unintended results (4). This process is based on continuous careful monitoring of the course/implementation of the programme.

Like in other fields, also on the field of education, we meet programmes. In this case we are talking about educational programmes. As mentioned at the very beginning of this module, in tight relation to this term is another term, being "curriculum".

Educational programme (program) versus curriculum

According to Licen (2), today both terms are used in planning of adult education process but while the term "program" is more frequently used in American literature, the second is more frequently used in British one.

1. **Educational programme**

   The term "educational programme" has several meanings, depending on context (2). It could designate several elements, a group of elements, or a whole system of elements:
   - in the case that it designates a plan for educational process, it means a design and documentation for implementation of educational process,
   - in other circumstances it means all educational activities, planned in selected environment (e.g. a school programme, a programme of an educational centre, etc.),
   - in a narrow meaning, it designates a sequence of teaching units, planned to reach selected goals in a given time. In this case, Jaksic is rather talking about "teaching plan" (1).

2. **Curriculum**

   Basically, the term "curriculum" originates from Latin word "currere" that means "to run". Accordingly, "curriculum" means "the way the process runs" or "a course" (2,7).

   In British didactics, with the term "curriculum" different learning and educational programmes prepared for participants of educational process are denoted (2). In the past, according to Jarvis (8), the concept of

curriculum was typical for elementary education, while in adult education the term "programme" was used. This terminology tried to make distinction between children and youth education, and adult education concept. Additionally, the term "programme" characterized more non-formal way of education, which was used in adult education process (8,9).

In andragogy today, the term "curriculum" is becoming more and more a general term for all educational programmes, irrespective of being more or less structured, and irrespective of being "opened" (usually in andragogy) or "closed" (usually in pedagogy) (2,8). Educational programmes for adults are not only non-formal. In fact, they are put in a continuum. There exist a pleiade of different programmes, from very non-formal to very structured, from very general to very specific. On the structured part of this continuum are programmes which are very close to traditional curriculum (traditional institutionalized) (8).

We will give some explanations related to types of educational programmes later on. In fact, we could define a curriculum as organized agreement on educational process and its content, to fulfill predefined objectives (8), yet in the field of andragogy such a definition should be understood as flexible as possible.

In higher education, which is somewhere in between pedagogy and andragogy (10), this term is usually used to designate the frame document in which professional and political agreements and decisions on educational role of a higher education institution (1).

At the end we could conclude that both terms could be used for the purpose of this module, but we decided to use the term "educational programme" in the meaning of a design of educational process, and documentation for its implementation. The reason is that in public health, as stated in a module "Public health capacity building: adult education principles and methods" in this book, we are facing with both parts of the educational continuum of adults. In capacity building we have very structured programmes, very similar to traditional curricula, while in educating process of educators and lay people, we are facing non-formal education process. Since this module is not intended to get to know students with knowledge how to plan higher education curricula, but rather more general knowledge, in the second part of this module, a planning process of a non-formal education programme will be presented.

From the PH academic sphere point of view, the term "curriculum" is more specific and is more connected to formal education in PH in higher education (capacity building) while the term "educational programme" is more general.

Planning

Planning could be defined as attempting to shape and control events in the future.

We plan because we want to increase chances to achieve the preset goals and objectives as much as possible. "The days are now past when the teacher produced a curriculum like a magician produced a rabbit out of a hat and when the lecturer taught whatever attracted his or her interest. It is now accepted that careful planning is necessary if the programme of teaching and learning is to be successful." (11).

As in many other planning processes, also in educational programme or curriculum planning process general recommended set of steps to be followed exists, being used by different experts slightly differently. In Table 1 ten steps in curriculum planning proposed by Harden are presented (12).

Table 1. Ten steps in curriculum planning according to Harden (12)

Step
1. Identifying the need
2. Establishing the learning outcomes
3. What to learn: content
4. How to learn: educational strategy
5. How to learn: teaching and learning tools
6. Assessment
7. Educational environment
8. Communication about curriculum
9. Managing the curriculum

Educational programmes planning models

In adult education, planning of an educational programme is a key part of this process. There exist several models of planning (2,8,13). According to Caffarella (13) the "program planning models consist of ideas about how programs should be put together and what ingredients are necessary to ensure successful outcomes". They could be different in shapes and sizes. They could be very simple with steps 1-5 for example, or very complex, using highly developed flow charts or in-depth qualitative descriptions (13). Some models are linear, some are non-linear (8,13):

- in linear models, the planner is expected to start at step one and follow each step in order until the process is completed. This may be helpful to newcomers, but soon loses its appeal since it does not represent the day to day working reality of most programme planners;
- non-linear, dynamic or non-sequential model allows programme planners to address a number of the components simultaneously, and to rearrange components to suit the demands of different situations. In these

models of programme planning is seen as a process that consists of a set of interacting and dynamic elements or components.

Among the most well known models are those proposed by Houle (Houle's decision points approach), Knowles (Knowles' andragogical model of programme planning), and Caffarella (Caffarella's interactive planning model) (13-15). In health education also PRECEDE/PROCEED model is well known (16). The detailed discussion on these models is far beyond the scope of this module.

In summary, the educational programme needs to define the learning objectives and learning outcomes, the setting in which it should be performed and the standard to which it should be performed.

Educational programmes planning

Elements of an educational programme

Irrespective if we name a programme of an educational process "curriculum" or "educational programme", it has indispensable elements. If we assume that every educational process and every educational activity could be fully planned, then the most simple structure of an educational programme (according to Tyler) could be used, being (8):

- objectives
- content
- methods of teaching and learning, and
- evaluation.

In practice it is not so easy, since the list of indispensable elements depends on how a programme planner understands the concept of education, and the concept of organized educational process.

Table 2. The scope of educational programme planning (modified Leinster's proposals) (17)

The scope	Description
Content	What knowledge, skills and attitudes should the programme cover?
	What are the learning objectives and learning outcomes of the programme?
Delivery	How will the learning be delivered?
	What teaching or learning methods will be used?
Assessment	Will the students' learning be tested?
Structure	How will the content be organized?
	How will learning and teaching be scheduled?
Resources	What staff, learning materials, equipment and accommodation is needed?
Evaluation	How will the organizers know that the educational programme has been effective in delivering the learning objectives or learning outcomes?

In summary, Leinster describes the scope of the educational programme planning as presented in Table 2 (17).

Educational programme planning steps

Different planning models list different but yet similar steps in planning process. In this module we are presenting the general proposal (1,2,8):

1. **Analysis of the circumstances and potential participants**

   Since educational programme is to be planned for particular environment its influences should be known and considered at the very beginning of the educational programme planning process. Analysis of these influences is important for assessment of internal (i.e. tradition of the educational institution, prevailing patterns of communication, its mission, financial resources, manpower, equipment, etc.) and external factors (educational market, competing institutions, local community characteristics, etc.). Both could be favourable or unfavourable.

   Second very important task at this first step of the planning process is analysis of the structure of potential participants of the educational programme. The data include data on gender, age, education background, socio-economic situation, geographical distribution, etc. For programme planner it is particularly important to be familiar with the educational background and previous knowledge of their target audience, since programmes should be planned to meet the educational needs and levels of the participants:

   - basic level - for those who have no previous knowledge of the topic,
   - intermediate level - for those who have some experience and basic knowledge in the topic being addressed,
   - advanced level - for those who are experienced, knowledgeable of the topic being addressed.

2. **Needs assessment**

   At this step, the needs and interests of the potential programme participants are identified. They should be identified on the basis of data sources (i.e. surveys) rather than on the basis of programme planner's opinions. This kind of surveys usually include questions on educational level of the participants, topics of the most

interest to participants, speakers of most interest to participants, participants' preferences on time, date, and location of the programme, reasons that might keep a participant from attending (transportation problems, financial limitations etc.).

There is often a mismatch between what is expected and the competencies gained from the training programme. The relevance or appropriates of educational programmes are the most important in designing and planning.

According to Dunn and co-workers, different approaches could be used to identify the curriculum needs (18):

- consultation with the stakeholders;
- a study of errors in practice;
- critical-incident studies;
- task analysis;
- study of star performers, etc.

3. Educational programme goals, learning objectives, and learning outcomes

After the outline of the programme is developed, learning goals, objectives and outcomes should be identified.

Learning objectives and learning outcomes are connected to the domains of learning (Table 3).

Table 3. Cognitive processes in Bloom's revised taxonomy for knowledge and descriptors of outcomes (19,20)

Bloom's taxonomy	Meaning	Outcome descriptor
Remember	retrieving relevant material from long-term memory	recognise, recall
Understand	determining the meaning of instructional massages, including oral, written, and graphic communications	interpret, make examples, classify, summarise, infer, compare, explain
Apply	carrying out or using a procedure in a given situation	execute, implement
Analyze	breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose	differentiate, organise, develop/change attitude
Evaluate	making judgements based on criteria and standards	check, critique
Create	putting elements together to form a novel, coherent whole or make an original product	generate, plan, produce

Educational programme goals are broad statements that define what the programme is expected to achieve. There exist short- and long-term goals. Short-term goals could be easily stated in behavioural terms and could be accomplished in a short time while long-term goals are more general and usually could not be accomplished in a reasonable amount of time.

Learning objectives are the statements that map out the learner's and planner's tasks needed to reach the goal. They state the specific knowledge, attitudes and behaviour changes needed to achieve the goal. In Table 4 some examples of learning objectives are presented.

Table 4. Some examples of learning objectives

Learning objectives
Upon completion of the programme, the participants will (be able to): • have an understanding of … • define and describe the … • distinguish the different types of … • gain new knowledge of … • gain skills on … • become aware of …

4. Content

The primary focus of a curriculum is many times on what is to be taught and when, leaving to the teaching profession decisions as to how this should be done. Many planners of an educational programme believe that the starting point for constructing a curriculum lies in the formulation of content. Many of them think in terms of what content students should learn. The fact is that when starting from this point not taking into consideration other curriculum elements there could be problems. It is very important to employ objectives in some way when constructing an educational programme. For formulating objectives analysis of the circumstances and potential participants is of enormous importance. In practice, also, how a topic is to be taught often determines what will be taught.

5. **Teaching methods**

In parallel with determining the content, decision must be taken about how it will be delivered. While there is debate over which learning methods are the most effective and efficient, there are some principles that should be taken into account. It is well known that knowledge is applied most effectively when it is learn in the context in which it is to be applied. It is also accepted that active learning is more effective than passive learning.

In this step the means through which the changes to be made are planned. Teaching methods like lectures, lectures with discussion, demonstrations, seminars, videos, role playing, case studies, etc. could be used. These methods are extensively discussed in a module "Public health capacity building: adult education principles and methods" in this book.

6. **Financial plan and marketing**

Educational programmes will almost always involve expenses of varying degrees. Regardless the expenses are minimal or large, they should always be planned in advance. Expenses mainly include speaker expenses, meeting/lecture room expenses, meal expenses, educational material expenses, and marketing expenses.

Sources for covering expenses could be fees or other sources of budgeting. When fees are to be charged they should be large enough to cover the costs but small enough to be marketable to target audience. It is worth to look at the costs of similar programmes and set registration fees accordingly.

Marketing the programme is also important for successful realization of the educational programme. A brochure or flyer is one of the most useful promotional materials. All information about the programme (e.g. description of the course content, learning objectives, speakers' names and qualifications, level of the programme, prerequisite skills or knowledge, date, time and location, etc.) should be accurately described. In addition to brochures/flyers, also other methods of publicity could be used, being press releases, public announcements, posters, etc.

7. **Course of the programme**

Course of the programme include identification of the speakers and their recruitment, location of the educational programme and logistic.

The basic criteria for selecting appropriate speakers are their knowledge and experience in the field of content of the educational programme, and their ability to mediate this knowledge to other people. In the process of recruitment, a summary of the programme should be discussed, the target audience described, and the level of the programme clearly specified. Length of the session, teaching aids, and logistic of the programme should be clearly introduced to potential speakers.

The location for carrying out the educational programme should be carefully chosen. Criteria for selection should include characteristics like adequate premises, clean and safe environment, close public transportation or available parking etc. There are usually many choices of meeting facilities e.g. hotels/motels, congress centres, community centres, schools, etc. It is important that size of the lecture rooms match the size of the expected audience (it is difficult to concentrate in a crowded or half full room). Also the possibility for providing meals and/or refreshments should be taken into consideration.

8. **Evaluation**

The last step to be planned in educational programme planning process is evaluation of teaching. This step has many different meanings:

- finding out the value of teaching;
- systematic way of learning and using the lessons learned;
- complex process of assessment;
- measurement and judgement;
- getting the best out of what you do for all involved, collection of information etc.

In respect to evaluation of teaching work it is frequently emphasized the importance of:

- students' opinion about the quality of teaching and engagement of teachers;
- measures outcomes of teaching (final knowledge of students);
- analysis of curricula and teaching by Faculty administration;
- occasional visits of departments and discussions with teaching personnel, etc.

This evaluation of teaching is rarely or never done. Also, in the world documents which very much emphasize the quality improvement in teaching process, the methodology for achieving these goals is not precisely defined, until, perhaps, recently (21,22).

In students' opinion about the quality of teaching one of the most common methods is student's anonymous questionnaire. However, it is generally believed that students' evaluation about teaching contribute to quality of teaching, although it is thought that its methodology needs to be still improved (23).

Programme evaluation involves getting feedback from participants as to how the programme met their needs and expectations. It should be done immediately after the programme ends and be carried out by completing an evaluation form. Areas covered in the evaluation form usually include programme content, speaker

effectiveness, facilities, handouts and other materials. Nevertheless, the evaluation form should be short and easy to complete. It should also allow the participant to make comments or suggestions.

One of methods that could be used in teaching evaluation is so called SITE (from Structured Interactive Teaching Evaluation) method[28] (24,25).

Case study

Designing and planning educational programme "Environment protection – what could we do?"

Introduction to the Case Study

This Case Study is a virtual case study, basing on a seminar paper at Ljubljana University Faculty of Arts, Department for Andragogy and Pedagogy, subject General andragogy II (26). Its characteristics are as follows:

- the virtual plan of the educational programme with public health content was constructed upon recommendations for planning adult education programmes of Slovenian Institute for Adult Education Ljubljana University Faculty of Arts (1,8);
- consultations, related to the public health content were given at Ljubljana University Faculty of Medicine, Chair of Public Health.

For the purpose of this module, the seminar paper was slightly supplemented and changed by its authors. For example, the background of the planning process was set to a slightly changed scenario.

Assumptions, underlying the virtual case study

Underlying assumptions to the new scenario are as follows:

1. **Assumption 1**

   The Ministry for Health of the Republic of Slovenia in cooperation with the Ministry for Environment of the Republic of Slovenia in the frame of sustainable development announced a call for application for a grant in maximal amount of 2,500 € for educational programmes related to environmental health. One of tender fields was "Health, environment, and waste disposal".

   The background for this tender field is rapidly growing problem of waste disposal in Slovenia. We are facing with increasing quantity of waste while there are less and less suitable locations for waste deposits. This situation is tightly connected with problems of era of postmodernism.

   In postmodernism, consumption has developed to undreamed extensions. In people it raises values, which are directed towards individuals and into the present time, not taking care of community's interests, or the future. As a result, negative consequences for the mankind are already starting to show. The problem has grown to the point at which we are starting to think over sustainable development. Hopefully, it is not too late.

   According to the World Commission for Environment and Development (WCED), sustainable development is a kind of development, that suits today's needs, but at the same time, it doesn't jeopardize chances of future generations to satisfy their own needs. The strategy of sustainable development embraces three pillars: economic development, social development and protection of environment. We are aware, how important the protection of environment is, therefore we need educational programmes, that would include that topic.

2. **Assumption 2**

   The Chair for Public Health of Ljubljana University Medical Faculty decided to prepare a proposal for an educational programme and apply for a grant. To fulfill all necessary standards of good educational programme, it was decided to ask also experts in pedagogy/andragogy to participate.

The process of designing, and planning

Analysis of existing circumstances and potential participants

Analysis of existing circumstances

In the process of brainstorming where is most necessary to start educate people how to preserve our nature healthy for today and tomorrow, we came to the conclusion that personal body care means are becoming a real threat to the nature. Among them are disposable diapers, and manufactured washing agents.

1. Disposable diapers

[28] Structured Interactive Teaching Evaluation (SITE) is an innovative method in evaluation. Starting from the fact that the evaluation of the teaching process is humane and not a mere administrative relationship with students and teachers as the partners, the original method of a structured interactive teaching evaluation was developed at the Department for Educational Technology, Andrija Stampar School of Public Health, Medical School, University of Zagreb. The SITE method has several characteristics:

- the evaluation process does not comprise all students but only randomly selected ones in the number corresponding to the number of all teachers in a particular course,
- participants in the evaluation process (selected students and teachers) are formally appointed by the Dean,
- final interactive meeting through "face-to-face" conversation is facilitated by a person from "outside".

The study on effectiveness and efficiency of the SITE method was done, comparing the results obtained by a traditional anonymous student questionnaire. The SITE method produced better impact on preparedness and active participation of both students and teachers, and influenced their responsibility in the evaluation process. The structured conversation added awareness to the teaching process and coordination to the proposals how to reach quality assurance jointly as equal partners. The SITE method was also efficient in terms of workload, time and cost of application.

When searching for rationale we have found that on the Technical University in Graz, Austria, in 1997 the study about pollution according to diapering system was done (27). The study has shown amazing results: production of disposable diapers contributes at very high degree to the pollution while pollution due to transport and due to waste is comparable.

When using washable diapers, pollution is much smaller and washing contributes most to it. This factor can be decreased by choosing environment friendly washing agents and by dosing them wisely up to 50% less then indicated from the producers. Thus we decided to go this direction.

2. Disposable sanitary towels

Additionally, we decided to include in our consideration also a problem, to the certain extent similar to the first one – use of disposable sanitary towels women use at menstruation. Enormous quantities of waste are produced by using this kind of sanitary towels.

In relation to this problem, we have found out that there exist environment friendly substitute, being a menstrual or menses cup. In Canada in 1995 even a clinical trial was performed to determine whether menses cup is well tolerated by menstruating women (28). The results showed that menses cup may be an acceptable method for some women for coping with menstrual flow. The menstrual cup exists now for decades, but only today its use is maybe put on the agenda. It was developed in the 1930s and went on commercial sale at around the same time as the first tampon. But unfortunately at that time it was not considered proper for women to touch their vaginal area. The manufacturers of the tampon overcame this obstacle by providing disposable applicators that can still be found floating in our toilet bowls today. Additionally, the tampon was commercially more attractive, as customers need to repurchase month after month. In contrast menstrual cup manufacturers do not expect to make repeat sales. Sales of tampons soared and massive advertising campaigns were launched. Disposable feminine hygiene products soon became established necessities in modern western society.

Thus we decided to analyze the circumstances of how could we preserve nature healthy by using selected environment friendly personal body care means (Box 1).

Participants

Who would be the participants of such an educational programme? If we limit to diapers, young parents are obviously the end target group. But not only this population group should and could be target group. Maybe educating young parents first wouldn't be the best choice. To targeting this population group as widely as possible, professionals that are coming in contact with young parents, well educated in this respect, seems to be the first choice.

But who would be the most suitable to make young parents aware of fact, presented above? We came to the conclusion that the best choice would be medical doctors and nurses. They are namely those, who accompany the child from even before birth on. That is why we should start to educate them, because they would be the ones, who will pass their knowledge to young parents.

Additionally, when searching for information on existence of educational programmes already offered in Slovenia about the similar topic, we discovered that nongovernmental organization (NGO) Association Storklja (a Stork) already holds lectures for parents on washable diapers in Maribor and Ljubljana. Since we didn't want to compete with this NGO, but rather supplement its efforts, we decided to target previously mentioned population group of medical doctors and nurses.

So, finally we decided to organise this educational programme for health workers, where they will be able to get needed knowledge about alternative hygienic accessories and skills, how to pass that gained knowledge to young parents (Box 1).

Additional rationale behind this decision is, that these professionals will be on the frontline not only in educating young parents how to behave friendly to environment, but also they should be prepared to answer numerous questions of end target group, etc.

Needs assessment

The next step is assessment of needs for an educational programme for educating people about harmful consequences of using environment unfriendly body care means and how to contribute to lessen this problem. We concentrated above all to disposable diapers usage problem.

We knew that sources of information on using disposable diapers are in Slovenia very limited, but we have found out, that NGO Association Storklja is leading a project entitled "Preserving the Water by Using Diapers Friendly for Earth and Water" (29). This project was granted from Regional Center for Environment for Slovenia and United Nations Development Program (UNDP) and Global Environment Facility (GEF) Danube Project (29,31). It is partnered by Association RODA from Croatia (29,32). The project emphasizes awareness-raising, environmental education, policy issues and public participation in decision making related to specific pollution problem of using disposable diapers. According to Association Storklja, in Slovenia, 99% of the population that uses diapers is using diapers that are one time use only, not realising how disputable they are from ecological point of view.

We could conclude that need for educating young parents of harmful consequences of using disposable diapers in Slovenia is obvious, and the problem should be tackled wisely (Box 2).

Box 1. The plan of educational programme "Environment protection – *what could we do?*" - Section 1: Analysis of the circumstances and potential participants

1. ANALYSIS OF THE CIRCUMSTANCES AND POTENTIAL PARTICIPANTS

1.1. ANALYSIS OF THE CIRCUMSTANCES

Body care means are becoming a real threat to the nature, since they represent considerable part of environmental pollution. Among them are also disposable diapers, manufactured washing agents, and disposable sanitary towels.

Disposable and eco-diapers

If using disposable diapers in a diapering period, the baby needs approximately 5,000-6,000 disposable diapers, and consequently contributes about 1 tone of waste. It is estimated that disposable diapers today represent about 8-10% of waste in industrialized countries, contributed by only about 1% of population. Disposable diapers seem at the first sight harmless for environment, since they are made of paper. But unfortunately this is far from the truth. To make them ultra-absorbing very dangerous chemical components are used: sodium poly-acrylate (ultra-absorbing gel), toxic organo-chloridic combinations and polyvinyl-chloride (PVC) that makes disposable diapers impermeable for wetness. This hazardous substances cause water pollution and it is suspected that they have toxic and carcinogenic or bio-accumulative effects, in particular because the they decompose very slowly - when deposited disposable diaper needs 500 years for decompose (29). Additionally, they are a source of greenhouse gases since babies' excrements are thrown to waste deposit in the disposable diaper and decompose there. Also many micro organisms can be found in thesa diapers. They slowly trickle trough toward the underground waters and can possibly infect the drinking one.

On the other side, the use of cotton diapers or other natural materials represents a great contribution towards environment protection, and what is even more, each and every family can take a part in it, if they only decide for environment-friendly way of changing their babies. If we use natural materials, the excrements in diapers are being removed into sewage system. Besides that each baby needs only 20-25 cotton diapers that could also be used for other babies. Special cotton diapers are eco-diapers. These are even more environment friendly. Eco-diapers are washable diapers, made of eco-cotton. Eco-cotton is cotton that grew without being sprayed with insecticides and herbicides and without chemical supplements. It had been harvested without any artificial additives. That kind of cotton is the most natural material and it pampers baby's skin as softly as disposable diapers (29).

Environment-friendly washing agents

Use of eco-diapers is related to another environmental problem – the use of washing agents. But pollution when using washable diapers is much smaller and washing contributes most to it. This factor can be decreased by dosing manufactured chemical washing agents wisely up to 50% less then indicated from the producers, or by choosing environment friendly washing agents.

Environment-friendly washing agents are detergents, made on soap-basis. Into that group we could place all products, that are soap-based, and all natural washing agents, for example washing nuts, soda, vinegar or ashes (29).

Menses cup

Enormous quantities of waste disposal are produced by using disposable sanitary towels, women use at menstruation as well.

There exists environment friendly substitute, known as menstrual or menses cup. The menses or menstrual cup is a reusable menstrual cup around two inches long and made from soft silicone rubber. It is worn internally like a tampon but collects menstrual fluid rather than absorbing. Unlike tampons the menses cup is not a disposable product, so only one is needed (30).

1.2. ANALYSIS OF POTENTIAL PARTICIPANTS

From public health point of view two population groups could potentially be the most suitable participants of this educational programme. The first group, that is also an end target group, is the group of young parents, or young future parents. The second group is consisted of health workers that are in contact with young parents, being medical doctors like obstetricians and paediatricians, midwives, and nurses like nurses in the community health centres, as well as home care nurses.

We decided to organise this educational programme for health workers, where they will be able to get needed knowledge about alternative hygienic accessories and skills, how to pass that gained knowledge to young parents.

Optimally, at least one medical doctor and one nurse from each participating health institution should participate at the seminar.

Box 2. The plan of educational programme "Environment protection – what could we do?" - Section 2: Needs Assessment

2. NEEDS ASSESSMENT

In Slovenia, 99% of the population of young parents is using diapers that are one time use only, not realising how disputable they are from ecological point of view. This leads to the estimate that in Slovenia every year about 45.000 children under age 2,5 years are using diapers what makes about 100 millions pieces of disposable diapers or 20.000 tons of waste that need to be deposited (29).

On the other hand, in Slovenia we are facing with problem of lack of suitable locations for waste disposal. People are refusing to have waste deposit in their neighbourhood, especially if the waste is produced in urban settlements, and are deposited in rural areas. These facts clearly indicate that there should be something done. Waste incineration seems to be one of possible solutions, but not a long-term one, since incinerators could also pollute the environment. Speaking in the medical terms, the best solution of all is to act etiologically, what means reduction in production of waste. For this could be realized, people need to be educated and enlightened.

Regarding problem of disposable diapers, education of young Slovene parents of harmfulness of disposal diapers on one hand, and use of washable diapers instead as an intelligent and rational choice, seems to be reasonable intervention. On Slovene market cotton-made and washable diapers could be purchased, but only a minor percent of young parents use them. The reason is, that one-time-use-only diapers are much more handy than washable ones. At the same time, parents don't realise, how much damage they make to the environment, when they are using disposable diapers instead of washable ones. That is why we should make them conscious of positive effects the use of washable diapers brings.

To be as much effective as possible, the problem should be tackled by educating directly young parents, and by educating health workers for they could educate young parents as well.

Learning objectives

We identified learning objectives at the macro and at the micro level as presented in Box 3.

Box 3. The plan of educational programme "Environment Protection – *What could we do?*" - Section 3: Learning Objectives

3. LEARNING OBJECTIVES

3.1 MACRO-LEVEL LEARNING OBJECTIVES

As the purpose of this programme the employees in community health centres and other health institutions should become aware of different ways, how to protect the environment, and of effects the environment has on people's health. Furthermore, we wish them to gain skills, how to pass to other people the knowledge they have got about how to protect the environment from pollution and about instruments to achieve that.

3.2 MICRO-LEVEL LEARNING OBJECTIVES

The participants:

- will become aware of the consequences of the environmental pollution in general,
- will become aware of the health impact of the environmental pollution,
- will gain new knowledge about positive effects of the use of environment-friendly hygienic accessories, from the point of view of different professions,
- will gain skills, how to pass their knowledge about environment-friendly hygienic accessories to other people, especially to young parents, and how to use them.

Box 4. The plan of educational programme "Environment protection – what coul*d we do?*" - Section 4: Contents

4. CONTENTS

Introductory lecture on ecology

An expert from the field of ecology is going to present negative consequences of non-ecological behaviour. Everything will be illustrated.

After the lecture, selected sections of the documentary film "Unpleasant truth" will be presented.

Lecture on ecological problem of disposal hygienic accessories, especially disposal diapers

In continuation, an expert from the field of environmental health is going to give a special attention to the problem of environment pollution caused by disposal diapers and disposal sanitary towels. Results of different studies will be presented, as well as suggestions how to cope with this problem.

Is this problem concerning health workers?

Health workers are mainly directed in curing a disease and similar, while prevention is not yet in the first plan, especially when is related to community health rather than to a health of an individual. Yet health workers are and should be those experts people have confidence in, and their advices and ideas still count in the community. They will be acquainted with the concept of empowerment of lay people how to behave healthy to the nature and the community.

Special attention will be given to that, how health workers in community health centres, working with young parents and future parents, and maternities, can do something, what would contribute to acting more ecologically in relation to the ecological problem of disposable diapers.

Lectures of different experts from several fields of activity about positive effects, that come as a result of using cotton diapers

Experts from different spheres are going to tell their points of view on use of washable cotton diapers instead of disposable ones and present their opinions. In such a manner, the participants will be able to realise, that the use of eco-diapers has desirable effects in the field of protecting the environment as well as in the specialized medical fields like dermatology and paediatrics.

The participants will watch A DVD film, entitled "Environment - and Earth-friendly diapers", that was prepared in 2006 by nongovernmental (NGO) Association "Storklja".

Representatives of the same NGO will demonstrate cotton-made diapers. Sellers of cotton diapers will exhibit these products at the exposition for participants of the programme could come in touch with them.

Lecture on natural washing agents

There is going to be a lecture on what natural washing agents are, and on positive effects of the use of them. Different experts, who collaborate with the Association "Storklja", will present results of some studies.

Lecture on menses cups

There will be a lecture on what a menses or a menstrual cup is, how it works, and its history. A results of a clinical trial will be presented.

Representatives of the Association "Storklja" will demonstrate an example of a contemporary menses cup.

Work in groups

In small groups, participants will think about how to transmit gained knowledge to other people, especially young and future parents. Each group will, with a help of a facilitator, prepare an assignment. Every assignment will be about finding new ways, how to promote the environment friendly products like eco-diapers, natural washing agents, and menstrual cups.

Contents

While modelling components of the contents of proposed educational programme, we tried hard to form such a programme that would help the participants to gain a total knowledge (Box 4).

Teaching methods

Teaching methods should be appropriate for the target group – young and middle-aged adults with college or university medical/health sciences background. We considered this and andragogical concepts of educational process were applied (Box 5).

Box 5. The plan of educational programme "Environment protection – *what could we do?*" - Section 5: Teaching Methods

5. TEACHING METHODS

The following teaching methods will be used:

- lectures will be performed for introduction to each session,
- demonstrations of environment friendly hygienic means will be uses whenever possible,
- buzz groups after introductory lectures,
- reflections and discussions in large groups,
- brainstorming, while working in small groups,
- discussions in small groups,
- plenary discussions, etc.

The educational programme will be organized as a seminar. The seminar is going to take place at week-end.

Budget determination and marketing

From the organizational point of view, we considered financial plan of the programme as well as marketing (Box 6).

Considering budgeting, we assumed that the programme will be granted with 2,500 €. "House of Nature" Company, a company providing environment friendly products, and NGO Association "Storklja", will co-finance the programme. Also, each participant will be charged a minor seminar fee in the sum of 25 € (Box 6). All incomes and expenses were calculated on 60 participants.

Course of the programme

First thing regarding the organizational part of the plan was to identify appropriate lecturers. While modelling components of the lectures, we tried hard to form a programme that would help the participants to gain a comprehensive knowledge. That is why we have asked experts from different fields of activity to share their knowledge. We thus planned to invite experts from different fields of public health being ecology, environmental medicine, and social medicine on one side, and experts from different fields of clinical medicine, being paediatrics, obstetrics, gynaecology, and dermatology on the other.

We tried to assess how many participants would attend such a seminar. We wanted to be realistic and set the number of participants to 60. However, there are about 70 community centres in Slovenia at the moment, and 14 maternities. Thus, the number could be even bigger.

Box 6. The plan of educational programme "Environment protection - *what could we do?*" - Section 6: Financial Plan and Marketing

6. FINANCIAL PLAN AND MARKETING

6.1 FINANCIAL PLAN

Incomes			
	Grant		2,500 €
	House of Nature Company	1,000 €	
	Association Storklja	1,000 €	
	Seminar fees		1,500 €
	Total		**6,000 €**
Expenses			
	Administration costs	200 €	
	Material costs		200 €
	Lecturers' fees		900 €
	Accommodation costs for lecturers	1,200 €	
	Rent for the congress room	250 €	
	Use of LCD projector (11€ per hour)	135 €	
	Development of the programme 500 €		
	Beverages	200 €	
	Coffee breaks		770 €
	Lunch		1,400 €
	Total		**5,755 €**

6.2 MARKETING

Chair of public health of Ljubljana University Medical Faculty will send the invitation letter to join the seminar to every community health centre, and every maternity. At the same time, it will be recommended at least one medical doctor and one nurse should take part at it.

We planned to invite also the Association "Storklja" to participate in education of educators of young parents as well.

Regarding the location of the seminar, we wanted first to locate it on an eco-farm, but unfortunately no such a farm offers a room, where the lectures could take place. That is the reason we have decided the seminar should take place in business centre Ajda (complex of Terme 3000, Prekmurje). This congress centre is very well known and location seemed to be very appropriate, because most of participants live in that area, so they should not have any problems with transport. That location is in the region, the education is meant for and beside that, Spa and congress centre of Moravske toplice is getting bigger and more significant all the time.

The planned course of the educational programme is presented in Box 7.

Box 7. The plan of educational programme "Environment protection - *what could we do?"* - Section 7: Financial Plan and Marketing

7. COURSE OF THE PROGRAMME
7.1 ORGANIZATIONAL ASPECTS
Number of participants: 60
Programme performers:
Coordinators: NS, expert of adult education
AK, expert of adult education
Lecturers: MF, Dipl.Ing.Chem., specialist in ecology
IE, MD PhD, specialist in environmental medicine
LK, MD PhD, specialist in epidemiology
PN, MD, PhD, specialist in paediatrics
ZN, MD, PhD, specialist in gynaecology
AM, MD, PhD, specialist in dermatology
Representatives of the Association "Storklja"
Representatives of the House of Nature Company
Duration: 2 days
Location: Business Centre Ajda, (Terme 3000), Moravske toplice, Slovenia

7.2 REALIZATION
The seminar will last for 2 days, being Friday and Saturday. Lectures will be held by experts in ecology, environmental medicine, epidemiology, paediatrics, obstetrics and gynaecology, and dermatology. Representatives of House of Nature Company and Association "Storklja" are also going to take part at it. In continuation the agenda of the lectures is presented.
The lectures will be variegated by picture material and short film.
Constituent element of the seminar will also be the reflexions of the participants – they shall have the chance to express their opinion.
The participants will have the chance, to look at the products that will be presented during the seminar.
On Saturday, the participants are going to work in small groups. Each group will have to brainstorm about, what would be the best way to promote natural products and to convince young parents to use them.
At the end of the seminar the participants shall present their reflexions, and discuss about it.

7.3 SEMINAR AGENDA
FRIDAY
9.00- 9:30 Registration of participants
9:30-10:00 Opening of the seminar and welcome to the participants
Coordination: NS, AK
Session 1 WE SHOULD BE WORRIED ABOUT OUR ENVIRONMENT
Coordination: NS, AK
10:00-10:30 Introductory lecture on ecology (lecturer MF)
10:30-11:00 Selected sections of the documentary film "Unpleasant truth"
11:00-11:30 Coffee break
11:30-11:50 Ecological problem of disposal diapers (lecturer LK)
11:50-12:10 Is this problem concerning health workers? (lecturer IE)
12:10-12:30 Reflexions of the participants
12:30-14:00 Lunch

Session 2 COTTON DIAPERS
Coordination: NS, AK
14:00-15:00 Presentation of eco-diapers, short film – Storklja, House of Nature
15:00-15:30 Washing agents, a disadvantage of cotton diapers (lecturer LK)
15:30-16:00 Coffee break
16:00-16:15 Ecologist and his point of view (lecturer MF)
16:15-16:30 Environmental health expert and his point of view (lecturer IE)
16:30-16:45 Paediatrician and his/her point of view (lecturer PN)
16:45-17:00 Dermatologist and his/her point of view (lecturer AM)

17:00-17:45	Reflexions of the participants (coordination AK, NS)
SATURDAY	
Session 3 NATURAL WASHING AGENTS, MENSES CUP	
	Coordination: NS, AK
9.00- 9:30	Types and use of natural washing agents (lecturer MF)
9:30- 9:45	Environmental health expert and his/her point of view (lecturer IE)
9:45-10:00	Dermatologist and his/her point of view (lecturer AM)
10:00-10:30	Reflexions of the participants (coordination AK, NS)
10:30-11:00	Coffee break
11:00-11:15	Menses cup and its history (lecturer LK)
11:15-11:45	Presentation of menses cup and its use – Storklja, House of Nature
11:45-12:00	Gynaecologist and his/her point of view (lecturer ZN)
12:00-12:30	Reflexions of the participants (coordination AK, NS)
12:30-14:00	Lunch
14:00-15:30	Work in groups (coordination AK, NS)
15:30-16:00	Coffee break
16:00-16:45	Reports of the groups (coordination AK, NS)
16:45-17:30	Reflexions of the participants - all participants and lecturers
17:30-17:45	Closure of the seminar

Evaluation

At the end we considered evaluation of the educational programme through targeting all three parts involved: health workers, users of eco-products, and companies that sell discussed product (Box 8).

Conclusion

Planning, designing, and then implementation of such an educational programme is a challenge, since it is an opposite to all commercials that are just about how practical and how much one product costs. What is sad, nowadays, our society prefers to buy products that don't need any maintenance - it is certainly easier, if we can, after we have used something, just throw it away. We are literally being pushed into that sort of acting if we want to do something for our children and grandchildren.

Box 8. The plan of educational programme "Environment protection - what could *we do?*" - Section 8: Evaluation

8. EVALUATION

The educational programme will be evaluated through reflexions of participants. We will ask the participants kindly to share their opinions about what has been told during the lectures. In the end of the seminar, a plenary discussion shall follow, where they will be able to express their opinion on the whole happening. We will deal out some sheets of paper and the participants should write down their reflexion. Like this, we will get the information about the quality of our programme (its contents as well as its organization).

Six months after the seminar will take place, we will make a research, if people are more interested in the products that have been presented in the seminar. We will ask every community health centre, if they could write a report on promotion activities that refer to eco-products. We are also going to make an opinion poll and ask parents of babies and women:

- if they have noticed any promotion of that products,
- if they maybe use one of them themselves, and
- what is their opinion on these products.

We plan to apply to the companies that sell discussed products to write a report on sales. We will ask them, if they think, making doctors known with these products has helped to increase people's interests in buying them

Acknowledgement

This sample plan was written under mentorship of Prof. Nives Licen, PhD, University of Ljubljana, Faculty of Arts.

Exercises

Groups of two, at maximum three students are formed.

Task 1

After introductory lecture, students carefully read the part on theoretical background of this module and corresponding recommended readings.

Task 2

As a second task they develop a teaching module. After this exercise, participants will be able to prepare and present their own module (learning objective).

In a process of development of a teaching module, students should pursue with the following teaching module structure:

1. Title
2. Author
3. Keywords
4. Rationale (educational needs)
5. Participants/students (to whom)
6. Duration
7. Learning objectives
8. Teaching methods
9. Planning of implementation
10. Assessment of participants/students
11. Module evaluation
12. Recommended readings
13. References

Presentation of a teaching module should be done in 30 minutes and assessed by other colleagues by using a questionnaire sample of which is presented in Table 5.

Table 5. An example of a questionnaire for an assessment of a teaching module

	Question	Not at all				Very much
1.	Module appropriate for target participants/students	1	2	3	4	5
2.	Learning objectives clear/well defined	1	2	3	4	5
3.	Learning objectives achievable in planned time	1	2	3	4	5
4.	Teaching methods appropriate	1	2	3	4	5
5.	Implementation plan well presented	1	2	3	4	5
6.	Methods of student assessment appropriate	1	2	3	4	5
7.	Evaluation of the module well defined	1	2	3	4	5
8.	Innovative module	1	2	3	4	5
		Poor				**Excellent**
9.	Overall assessment of the module	1	2	3	4	5
10.	Oral presentation	1	2	3	4	5
Comments and recommendations to author (please describe in few words):						

References

1. Jaksic Z. Curriculum and planning of lessons (in Croatian). In: Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005. p.114-119.
2. Licen N. Introduction to adult education (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za pedagogiko in andragogiko; 2006.
3. Last JM. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
4. World Health Organization, Regional Office for Europe. Handbook for process evaluation in noncommunicable disease prevention. CINDI 2000. Copenhagen: World Health Organization, Regional Office for Europe; 1999.
5. Planning and Performance Management Unit. Office of the Director of U.S. Foreign Assistance. Glossary of Evaluation Terms. Washington D.C.: Planning and Performance Management Unit. Office of the Director of U.S. Foreign Assistance, 2009.
6. Swiss Federal Office of Public Health. Evaluation. Glossary of evaluation terms. Berne: Swiss Federal Office of Public Health; 2005. Available from URL: http://www.bag.admin.ch/evaluation/02357/02603/index.html?lang=en&download=M3wBPgDB/8ull6Du36WenojQ1NTTjaXZnqWfVp3Uhmfhnapmmc7Zi6rZnqCkkIR2fXqDbKbXrZ6lhuDZz8mMps2gpKfo. Accessed June 4, 2013.

7. The online etymology dictionary [homepage on the Internet]. Available from URL: http://www.etymonline.com/index.php. Accessed June 4, 2013.
8. Svetina M. Educational programmes for adults from planning to carrying out (in Slovene) Ljubljana: Andragoski center Republike Slovenije; 1998.
9. Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.101-221.
10. Jaksic Z. Student-centred teaching (in Croatian). In: Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005. p.87-95.
11. Harden RM. Curriculum planning and development. In: Dent JA, Harden RM, editors. A practical guide for medical education. London: Elsevier; 2005. p.10-8.
12. Harden RM. Approaches to curriculum planning. Medical Education. 1986;20:458-66.
13. Caffarella RS. Planning programs for adult learners: a practical guide for educators, trainers, and staff developers. San Francisco, CA: Jossey-Bass Inc.; 1994.
14. Smith MK. Malcolm Knowles, informal adult education, self-direction and anadragogy. The encyclopedia of informal education. London: Infed; 2002. Available from URL: http://www.infed.org/thinkers/et-knowl.htm. Accessed June 4, 2013.
15. Warren RM. Program planning and development in adult education: Where we are at the beginning of the 21st Century. Muncie, IN: Ball State University. Available from URL: www.hiceducation.org/Edu_Proceedings/Ruth%20M.%20Warren.pdf. Accessed June 4, 2013.
16. Campbell C. Health education planning models - a review of the literature - part II. Starkville, MS: Mississippi State University; 2001. Available from URL: http://msucares.com/health/health/appa2.htm. Accessed: June 4, 2013.
17. Leinster S. The undergraduate curriculum. In: Dent JA, Harden RM. A practical guide for medical teacher. London: Elsevier; 2005. p.19-27.
18. Dunn WR, Hamilton DD, Harden RM. Techniques of identifying competences needed by doctors. Medical Teacher. 1985;7:15-25.
19. Quality Assurance Agency for Higher Education [homepage on the Internet]. Available from URL: http://www.qaa.ac.uk. Accessed June 4, 2013.
20. Anderson LW, Krathowohl DR, editors. A taxonomy for teaching, learning and assessing: a revision of bloom's taxonomy for educational objectives. New York, London: Longman, Addison Wesley; 2001.
21. Guilbert JJ. Educational handbook for health personnel. Geneva: WHO; 1987. Available from URL: http://whqlibdoc.who.int/offset/WHO_OFFSET_35.pdf. Accessed June 4, 2013.
22. Standards and guidelines for quality assurance in the European higher education area. Helsinki: European Association for Quality Assurance in Higher Education; 2005. Available from URL: www.enqa.eu/files/BergenReport210205.pdf. Accessed June 4, 2013.
23. Metcalfe DH, Matharu M. Students' perception of good and bad teaching: report of a critical incident study. Med Educ. 1995;29:193-7.
24. Pavlekovic G, Bradamante Z. Structured Interactive Teaching Evaluation (SITE): an innovative method in evaluation. In: Association for Medical Education in Europe. Book of abstracts. Cotone Congressi Genoa, Italy, 14-18 September 2006. Dundee: Association for Medical Education in Europe; 2006. p.90.
25. Pavlekovic G. Evaluation of quality of teaching: why SITE (in Croatian)? In: Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005. p.299-305.
26. Skerget N, Kragelj A. Educational programme: Environment protection what could be done by ourselves. Subject: General andragogy II. A seminar paper. Ljubljana: University of Ljubljana, Faculty of Arts; 2008.
27. Association Storklja. Preserving the water by using diapers friendly for earth and water [homepage on the Internet]. Available from URL: http://www.storklja.si/t3/index.php?id=1103#1954. Accessed June 4, 2013.
28. Cheng M, Kung R, Hannah M, Wilansky D, Shime J. Menses cup evaluation study. Fertil Steril 1995;64:661-3.
29. Association Storklja [homepage on the Internet]. Available from URL: http://www.storklja.si/. Accessed June 4, 2013.
30. Mooncup. Whai is it [homepage on the Internet]. Available from URL: http://www.mooncup.co.uk/menstrual_cup_whatisit.html. Accessed June 4, 2013.

31. Regional Environmental Center. Danube regional projects: Regional grants [homepage on the Internet]. Available from URL: http://www.rec.org/REC/programs/NGO_support/grants/RegionalDanubeGrants/ResultsRound2.html. Accessed August 24, 2013.
32. Association Roda [homepage on the Internet]. Available from URL: http://www.roda.hr/. Accessed: June 4, 2013.

Recommended readings

1. Caffarella RS. Planning programs for adult learners: a practical guide for educators, trainers, and staff developers. San Francisco, CA: Jossey-Bass Inc.; 1994.
2. Dent JA, Harden RM, editors. A practical guide for medical education. London: Elsevier; 2005.
3. Jaksic Z. Curriculum and planning of lessons. In: Jaksic Z, Pokrajac N, Smalcelj A, Vrcic-Keglevic M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko drustvo za medicinsku edukaciju; 2005.
4. Licen N. Introduction to adult education (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za pedagogiko in andragogiko; 2006.
5. Svetina M. Educational programmes for adults from planning to carrying out (in Slovene) Ljubljana: Andragoski center Republike Slovenije; 1998.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Public health planning: From recommendation to implementation**
Module: 2.41	ECTS (suggested): 0.75
Author(s), degrees, institution(s)	Valeriu Sava, MD, PhD Helmut Wenzel, M.A.S. Centre of Public Health and Management, Chisinau, Part-time lecturer at the University of Medicine and Pharmacy, School of Public Health. The second author is health economist.
Address for correspondence	Valeriu Sava 53, Dacia Blvd, ap.224 MD 2062, Chisinau, Moldova Tel: (373-22) 769704 Fax: (373-22) 664018 Email: mam@moldnet.md
Keywords	Efficient implementation of policy recommendations, rational decision-making, system analysis.
Learning objectives	After completing this module students and public health professionals should understand how: • an implementation of projects and solutions are done within complex social systems, with many interdependencies of interests and motives, when planning takes place at different levels and in different organisational structures of a country or region, • one has to distinguish between rejection or delay of a decision due to real feasibility hurdles or politically correct but disguised refusals. • to identify and analyse the rules of the game in the corresponding "fields of force" and to set up models for simulating qualitatively possible interventions.
Abstract	Every analysis of the reality of planning has to take into account that planning takes place at different levels and in different organizational structures of a country or region. Such a plurality of planning authorities and corresponding values can lead to conflicts and - in the worst case - to stalemate. It is also important to really understand the meaning of a possible rejection or delay of a decision. Therefore it is vital to understand the motives and interests of involved stakeholders, the connectivity of goals, the interrelationship of organizations. Furthermore, the likely outcomes of any suggested strategy have to be evaluated from the viewpoint of the stakeholders, qualitatively and where possible also quantitatively. This helps to understand which stakeholder would be supportive, indifferent or possibly against a plan. Social systems are very complex systems with many elements and processes. They are regulated by positive and negative feedback loops. The relatively high number of feedback loops that act with different speed, strength of impact and direction (supporting or weakening) makes it impossible to predict actions or reactions of a system when introducing new activities. Systematic description and simulation of such systems is needed.
Teaching methods	After the introductory lecture students will participate in nominal group technique in order to recognize and to rank the field in the quality of health care where organizational, managerial, or other improvements are necessary, such as waiting lists, admission policy, medical records keeping, patient's discharge procedure, administration of drugs, working in multidisciplinary teams, patient satisfaction, etc. Then students will work in small groups, divided according to country or working place, to discuss the possibilities for improvement in their own environment. The second exercise will be to discuss, within the country (or working place) small groups, the necessary procedure for development of national accreditation system. Teachers will advise students to follow existing models and experience and to highlight their advantages and obstacles in the case of application within the country of SEE region.
Specific recommendations for teachers	No specific recommendations.
Assessment of students	Multiple-choice questionnaire (MCQ), and case problem presentations.

PUBLIC HEALTH PLANNING: FROM RECOMMENDATION TO IMPLEMENTATION

Valeriu Sava, Helmut Wenzel

Theoretical background

Introduction

The previous chapters dealt with different problems, aspects of uncertainty and tools for creating a scientific basis for analyses and policy recommendations. Besides this rational, scientific problem solving, understanding the peculiarities of the practical political problem solving process - i.e. politics - makes the difference between successful implementation and failure. As Reich states "Good politicians have an intuitive sense about the strategies that are likely to work in a particular situation, based on their accumulated experience with political life" (1). He further points out that most of the public health experts are not well trained in political analysis. More often, they believe that finding the right technical or scientific answer in their professional field of activity e.g. epidemiology or economics is enough (1). This scientific based perception finds its pendant in the definition of the corresponding professional role: the pure scientist (2). He accepts a division of labour between the scientific expert and the political professional; he only feels responsible for the proper handling of the issues. The implementation is not in his focus. A more appropriate course of action will not only bring the expert to take a wider view of his task, but also lead to the insight that certain skills are needed, to meet political decision makers on eye level. After all, the paradigm of the "honest broker" corresponds precisely to the portrayal made above. The "honest broker" is engaged in political procedures and processes, broadens the scope of research question in a way that he takes into account options of actions and its consequences from the viewpoint of the decision maker (2).

Political decisions are not primarily the outcomes of a scientific problem solving process as depicted in a rather normative policy cycle scheme. In fact, the decisions are made in an field of conflicting interests of politicians, voters, special interest groups, and finally have to reflect trade-offs (3). Policy analysts use a variety of analytic models (4) to describe and investigate the phenomenon. The spectrum ranges from a rather normative scientific problem solving approach to a more incrementalistic way (5) - Lindblom even calls it "muddling through" (6) - and finally to the paradigm of bounded rationality (7).

What applies to the political decision making process applies to the execution of the decisions as well. The implementation of the policy is far from being an easy task. Like in the political decision process, multiple actors are involved. They possibly will belong to different organisational entities and they may not have the same preferences concerning the way in which a given policy should be implemented. The specific interests of the various players are likely to affect the outcomes. Deviations in reaching policy goals are probable if the implementation process depends upon a number of players who have to cooperate. Therefore, the longer the implementation chain and the greater the number of actors involved in the process, the more difficult implementation becomes. If a given policy is introduced but insufficiently put into practice, it is possible that the final result will be less desirable than the status quo (7). As a rule, the problem that initiated the policy-making process can only be solved effectively if the adopted policy is properly implemented.

After scrutinizing the more technical aspects and the potential impacts for the various players, the political arena has to be analysed thoroughly. Because planning takes place in an organisational and political milieu, it is important to understand the "rules of the game" in the corresponding "fields of force". Planning occurs at different levels (micro and macro) and in different organisational structures of a country or region. Such a plurality of planning authorities with their corresponding values can lead to conflicts and - in the worst case - to stalemate. Many implementation problems seem to date from the fact that macro level implementers cannot influence micro level implementers (7). It is also important to understand the real reason of a possible rejection or delay of an implementation. It is essential to distinguish between real hurdles - such as scarce resources etc. - and politically correct but disguised refusals in the "garment" of specific circumstances, so called inherent necessities.

As a prerequisite for successful implementation, we need to understand how the various actors are involved, and how they cooperate, thus understanding varying motives and anticipating possible hurdles. Furthermore, the likely outcomes of any suggested strategy have to be evaluated from the viewpoint of the stakeholders, qualitatively and where possible also quantitatively. This helps to understand which stakeholder would be supportive, indifferent or possibly against a plan. Ideally, an expert would consider this when preparing his recommendations.

In this context, he ought to:

- Determine the likelihood to reach that goal with a given amount of resources;
- Anticipate and identify undesired effects that might occur due to the goal to be reached. This has to do with the complexity of the system where the activities take place;

- Balance the advantages of reaching the goal with the disadvantages of side effects;
- Analyse the opportunity cost with regard to forgone possibilities when using the resources for other purposes;
- Thus, evaluate the relative importance and relevance of the goal with respect to other values.

System thinking and system analysis

In Western cultures, thinking in linear cause-effect-relationships is very common. In combination with the fact that social systems - the arena of decision-making - are very complex, this fosters simplified views of "reality". In fact, our reality is determined by the concurrence of elements. Influencing factors are linked and evolve dynamically over time. Therefore, we need to think in terms of cycles and interrelations (feedback loops). This enables the analyst to follow the dynamics of a system and to identify possible unintended side effects and feedback effects. Very often, surprisingly very weak feedback loops, easy to overlook, can have a stronger influence on the course of a process than apparent direct interdependencies.

In this context, the "nested thinking approach" (8) was promoted. This means that analysts understand problems as a dynamic network of alternating influencing factors. Including all relevant relations into the analysis is a precondition for appropriate problem solving and anticipating the likely outcomes of an intervention. In other words this kind of approach tries to describe or to understand social systems. They are regulated by positive and negative feedback loops. The relatively high number of feedback loops that act with different speed, strength of impact and direction (supporting or weakening) makes it impossible to predict actions or reactions of a system when introducing new activities. Systemic description and simulation of such systems is needed. More precisely, the analyses focus on the allocation of power, the spread of the various kinds of influence - the make-up of the chain of reaction, the feedback loops, their changes over time (the dynamic perspective of the system), and possibly existing archetypes.

To assess and visualise the stakeholders, their power and their relationships different methods can be used. Firstly, a very simple, but very helpful way is to draw an influence diagram. It is based on the system analysis approach. Professional software tools like GAMMA® or TOPSIM® can support the sketch of the network. In contrast to a quantitative modelling (e.g. systems dynamics), only the strength of the connections needs to be recorded. Secondly, a Social Network Analysis (SNA), which is a more sophisticated method. From a technical viewpoint it is applied Graph Theory (9). Practically it can build on the network as designed by system analysis tool. SNA offers many indices that can describe the system and the importance of actors in the network, then. Several software packages are available. Thirdly, a kind of Computer-Assisted Political Analysis-Tool - PolicyMaker®. The software can be applied to any policy problem that involves multiple players with diverging interests. These methods differ by the required skills and the effort in terms of time and cost.

Outline of a system analysis sequence

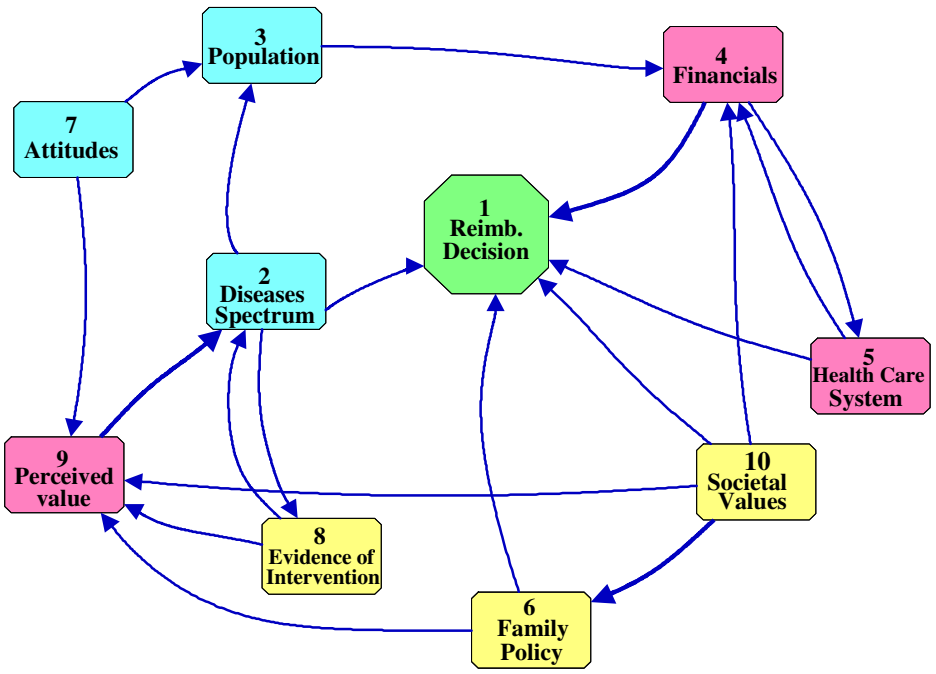
There exist varieties of recommendations and guidance tools how to identify and analyse stakeholders (1,10-12). The system analysis sequence mentioned above does not require too much effort. Therefore, we will focus on these steps.

Ideally, this process should be done in working groups. This modelling approach is rather qualitative, so no algorithms are needed. The information used can be intuitive or - in the best case - be empirical. The engine of the model is transparent, the computational skills is basic arithmetic operations; no black box is used, no mystic expert knowledge is hidden somewhere. The expertise is within the group, the members are the experts.

1. **Definition of the system** to be analyzed. Here we describe the field of activities and define the goals of the analysis. All the relevant elements and processes of the system under survey have to be listed. Exclusion of all elements that have not to be considered is a cross-check of that step and helps to define the boundaries of the system (see example).
2. **Modelling phase**. The elements of the system will be connected by lines symbolising the influence by strength, direction (causality) and time aspects. Very often those lines can be understood as processes. As a result of this step the complexity and the interdependencies are shown. This visualization is the basis for consensus building in the group: Do we have the same understanding of the boundaries, the elements and the processes of the system we should look at?
3. **Learning phase**. Basic analyses are carried out to understand the dynamics of the system. The simplest analytical step is a portfolio analysis dividing the elements according to the influence they receive or the influence they have on other elements into active and passive elements (see example). In a more sophisticated step the speed of influences and feedback loops are analysed. As prerequisite processes have to be evaluated by their level of strength and the speed of influence, (speed can be measured in terms of hours, days, or years – it depends on the system and the time horizon). In this case, it can be graphically shown how influences are evolving (see example). Again, this can be used also to build consensus on the dynamics and behaviour of the system under consideration.

4. **Strategy phase**. First, the possible options for interventions (or changes) have to be listed. Most preferably, this could be done alongside the model structure. This prevents overlooking of important options. In a next step, the consequences of interventions are simulated. Based on the learning the desirable activities are listed, maybe resources needed are estimated and responsibilities for the actions are agreed upon.

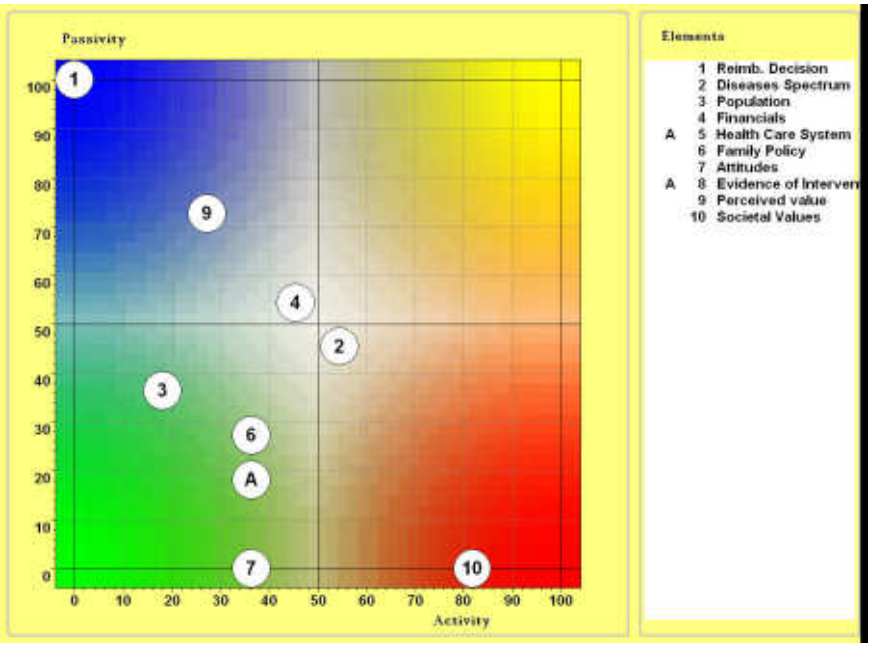
Figure 1: Influence diagram of a hypothetical network



Source: Based on H. Wenzel, Elements that influence Reimbursement Policies, presentation given at the Annual ISPOR Meeting in Barcelona, 2002

In this hypothetical case an intervention should be introduced, and reimbursed. The intervention is targeting at improving women's health with special focus on family planning. The health care system and the financial situation are constraints that will change from country to country. Aside from the technical (rational) aspect, like the evidence of the intervention in question, there are also elements that characterise the society, its inherent value system, and the corresponding family policy.

Figure 2: Portfolio of importance and influence



The actors are located according to their degree of activity and passivity - this matches the concept of outdegree and indegree of SNA. The actor's importance can now be interpreted based on the positioning (13). The quadrant on the lower right side shows the elements with the highest influence within the system: The value system of the society (#10) and the spectrum of diseases (#2). They are targets for efficient actions. Elements in the upper left quadrant have no or nearly no influence (#1, #9, #4). Any measures taken here to change the system is inefficient. Elements in lower quadrant on the left side (#3, #5, #6, #8) are inert. They are behaving like buffers and are inefficient for actions. Elements in upper right quadrant are critical, mostly due to feedback loops. This makes their influence and activities unpredictable.

Social Network Analysis primarily measures centrality. Centrality measures identify the most prominent players, i.e. those players who are extensively involved in relationships with other network members. The concept of centrality helps to identify key players (14). The most frequently used centrality measures are degree centrality, betweenness centrality and closeness centrality. They are based on the assumption that information is transferred along the shortest pathways. Scott (9) gives an extensive overview on the various measures and the different options for data collection. SNA has been applied in various situations. In Nigeria the introduction of new vaccines into national immunization programmes was analysed (15). Pagliccia et al. (16) used SNA to put to the test intersectoral cooperation of policy-makers in Cuba. Wenzel et al (17) developed a network of the Serbian Health Care System. Glasmann et al. (18) examined the major political challenges that were associated with the adoption of health reform proposals in the Dominican Republic.

A kind of Computer-Assisted Political Analysis (CAPA) (19)

PolicyMaker is a rapid assessment method for analyzing and managing the political decision making process. PolicyMaker is a logical and formal procedure to provide practical advice on how to manage the political aspects of public policy. The method helps decision-makers improve the political feasibility of their policy.

It is carried out through five main analytic steps:

- Policy Content: Define and analyze the content of the policy. Identify the major goals of the policy, and specify a mechanism that is intended to achieve each goal. Determine whether the goal is already on the agenda.
- Players: Identify the most important players and analyze their positions, power, and interests, and assess the policy's consequences for the players. Also, analyze the networks and coalitions among the players.
- Opportunities and Obstacles: Assess the opportunities and obstacles that affect the feasibility of your policy, by analyzing conditions within specific organizations and in the broader political environment.
- Strategies: Design strategies to improve the feasibility of your policy, by using expert advice provided in the program. Then, evaluate the strategies, and create alternative strategy packages as potential action plans.
- Impacts of Strategies: Estimate the impacts of your strategies on the positions, power, and number of mobilized players--the three factors that affect the feasibility of your policy. Compare the future and current Position Maps and Feasibility Graphs to show the impacts of your strategies. Monitor the implementation of your strategies and compare the results to your predictions.

The analysis in PolicyMaker results in a series of tables and maps or diagrams that systematically organise essential information about a policy. These tables and maps can be used in strategic planning for policy formulation and implementation, to assist in improving the political feasibility of a policy. The results can help with:

- Understanding - by facilitating the analysis of the political circumstances faced by a policy;
- Problem identification - by providing rapid identification and definition of obstacles;
- Policymaking process - by assisting in communication among different organizations;
- Organize data - by providing a database and easy-to-use screens to store, track, and analyze positions, power, and other aspects of a political question;
- Implementation strategies - by proposing new ideas and strategies, helping you to evaluate their consequences, and to track their implementation;
- Overall enhanced impact of your policy - by improving the chances that a policy will achieve its intended effects.

The concept

The software uses political mapping techniques to analyse the political actors in a policy environment. These techniques assess the power and position of key political actors, and then display the supporters, opponents, and non-mobilized players in a political "map" of the policy. This mapping provides the basis for designing strategies of political management. Furthermore, the software incorporates

techniques of political risk analysis, in order to provide a quantitative assessment of whether a policy is politically feasible. Political risk analysis methods have been adapted to assess the feasibility of a particular policy, through an algorithm that is calculated in several basic forms and can be modified by the user. PolicyMaker uses methods of organizational analysis and a rule-based decision system, in order to suggest strategies that can enhance a policy's feasibility.

Limitations

On the other hand, PolicyMaker is based on a series of assumptions about the policy-making process and a series of judgments about the characteristics of players and impacts of strategies. As with any analytical method, the validity of the PolicyMaker analysis is limited by the quality of these assumptions and the judgments of the analyst. The quality of the analyst can affect the reliability of the data, analysis, and interpretation, because this method involves subjective (but systematic) assessments of players, positions, and power, and speculation about the impacts of strategies. A second limitation is the analytical model contained in PolicyMaker. The software uses an algorithm to calculate an index of political feasibility, based on a formula that combines quantitative assessments of three factors (the power of players, the positions of players, and the number of mobilized players), as noted earlier. The feasibility algorithm can be modified in various ways to alter the model for assessing political feasibility. However, this model (as with all models) is still limited by its assumptions and simplifications about how the world works. The analysis depends on analyst's judgments about players and their power and positions, on his assessments of the impacts of specific strategies on players, and on the program's assumptions about interactions among strategies and among players.

References

1. Reich, MR. The politics of reforming health policies. IUHPE - Promotion & Education, 2002, IX/4:138-142.
2. Pielke, Roger A. The honest broker. Making sense of science in policy and politics. Cambridge, Eng, Cambridge University Press, 2007.
3. Schneider, V. and Janning, F. Politikfeldanalyse. Akteure, Diskurse und Netzwerke in der öffentlichen Politik. Wiesbaden, VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH, Wiesbaden, 2006.
4. Parag, Y. A System Perspective for Policy Analysis and Understanding: The Policy Process Networks. The Systemist, 2006,28:212-224.
5. Hayes, Michael T. Incrementalism and public policy. New York NY u.a, Longman, 1992.
6. Lindblom, C. E. The science of "muddling through". Urban management, 1972, 96-105.
7. Knill, Christoph and Tosun, Jale. Public policy. A new introduction. Basingstoke, Hampshire u.a., Palgrave Macmillan. 2012:2.
8. Probst, Gilbert and Gomez, Peter. Vernetztes Denken. Wiesbaden, Gabler, 1989.
9. Scott, J. Social network analysis: A handbook. 2. ed., reprinted ed. London u.a., Sage Publ, 2009.
10. Varvasovszky Z. and Brugha R. How to do (or not to do) ... A stakeholder Analysis. Health policy and planning, 2000,15:338-345.
11. Bryson John M. What to do when stakeholders matter - A Guide to Stakeholder Identification and Analysis Techniques. Public Management Review, 2004,6:22-53.
12. The World Bank. Tools for Institutional, Political, and Social Analysis of Policy Reform - A Sourcebook for Development Practitioners. Washington D.C., 2007.
13. Ulrich Hans and Probst, Gilbert JB. Anleitung zum ganzheitlichen Denken und Handeln - Ein Brevier für Führungskräfte. 2. ed. 1990:145.
14. Hawe, P., Webster, C., and Shiell, A. A glossary of terms for navigating the field of social network analysis. J Epidemiol Community Health, 2004,58:971-975.
15. Wonodi, C. B., Privor-Dumm, L., Aina, M., Pate, A. M., Reis, R., Gadhoke, P., and Levine, O. S. Using social network analysis to examine the decision-making process on new vaccine introduction in Nigeria. Health policy and planning, 2012.
16. Pagliccia N, Spiegel J, Alegret M et al. Network analysis as a tool to assess the intersectoral management of health determinants at the local level: A report from an exploratory study of two Cuban municipalities. Social Science&Medicine, 2010,71:394-399.
17. Wenzel, H., Bjegovic V., and Laaser, U. Social Network Analysis as a Tool to evaluate the balance of power according to the Serbian Health Insurance Act. Management in Health, 2011, 8-15.

18. Glassman, A., Reich, Laserson, K., and Rojas, F. Political analysis of health reform in the Dominican Republic. Health policy and planning, 1999,14:115-126.
19. Reich MR. Applied Political Analysis for Health Policy Reform. Current Issues in Public Health, 1996, 191-196.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Team buiding**
Module: 2.42	ECTS (suggested): 0.25
Author(s), degrees, institution(s)	Adriana Galan, Public Health Consultant National Institute of Public Health, Romania Silvia-Gabriela Scintee, MD, MSc, PhD, National School of Public Health, Management and Professional Development, Bucharest
Address for Correspondence	Name: Adriana Galan Institution: National Institute of Public Health Street: Dr. Leonte 1-3 City: Bucharest Country: Romania Tel: + 4021 3183620; Fax: + 40213123426 E-mail: adriana.galan@insp.gov.ro
Keywords	Formal/informal groups, group development, team building, team roles, team working.
Learning objectives	After completing this module students and public health professionals should: • Differentiate between group and team; • Recognize different types of groups; • Be aware of and identify possible roles within a team; • Understand the group development process.
Abstract	Because there are no pure formal or informal organizations in real world, one may conclude that an organization is a mix of formal and informal groups. We must differentiate the concept of group versus the concept of team. The literature describes several types of groups according to a set of criteria. Formal groups (work team) is created by an organization in order to achieve a certain goal, being recognised and receiving full support from the organization. Informal groups arise from natural attractions among individuals for social reinforcement or other benefits. The roles within a team actually represent tasks and functions in the self-management of the team's activities.
Teaching methods	Teaching methods include: lectures, group discussions, group assignments. First assignment consists in self-evaluation of the roles played in a team by each student, based on Belbin questionnaire. For the second assignment, the students will be split in groups of 4-5 persons and asked to solve the "Survival" exercise. Each group should evaluate the performance of the team in the process of decision-making versus the individual performance.
Specific recommendations for teachers	60% work under teacher supervision/ 40% individual students' work. Belbin questionnaire and Survival game should be distributed during the course.
Assessment of Students	Assessment could be based on multiple-choice questionnaire (MCQ), structured essay, seminar paper, case problem presentations, oral exam, etc.

TEAM BUILDING

Adriana Galan, Silvia Gabriela Scintee

Theroretical background

Generally, either formal or informal organizations constitute together the organizational setting where work is performed. As described by Organization Design theory, a formal organization is characterized by approved authority-responsibility relationships, clear division of work and separate departments and, usually, a hierarchical structure. The formal organization is the planned interrelationship of people, material resources and activities (1).

By contrast, the informal organization is characterized by dynamic behavior and activity patterns taking place within formal structures due to human interrelationship and interaction.

Because there are no pure formal or informal organizations in real world, one may conclude that an organization is a mix of formal and informal groups. Thus, its performance depends on the management ability to recognize the existence of these groups, to transform them from groups into working teams, to motivate and stimulate them to achieve organization's goals.

Nevertheless, to build an effective team is a resource consuming process and requires a complex activity.

Group versus team

We must differentiate from the very beginning the concept of group versus the concept of team. Thus, a simple definition of the group can be: two or more persons who come into contact for a purpose and who consider the contact meaningful (2). The purpose of a group may be implicit rather than stated. Another definition of a group is: two or more persons who communicate and share common values, norms and needs (3). To conclude, there are three minimal requirements to form a group:

1. group members are sharing common needs or objectives;
2. there is a strong interpersonal communication component within the group;
3. there is a minimal or even absent hierarchy within the group.

A team represents a small number of people with complementary skills who are committed to a common purpose, performance goals, and approach for which they are mutually accountable (4).

A team has to accomplish bigger goals than any individual group. The purpose of a team is to perform, achieve results and be successful in the organization or marketplace. A good manager is those who can assemble a group of individuals and transform them into a team. There were described ten key differentials to help a manager to shape a group of people into a pro-active and productive team (5).

1. **Understandings.** In a group (usually in formal groups), members think they are put together for administrative purposes only. Individuals sometimes cross purpose with others. In a team, members recognize their independence and understand that both personal and team goals are best accomplished with mutual support. Time is not wasted attempting personal gain at the expense of others.
2. **Ownership.** In a group, members rather tend to focus on themselves because they are not sufficiently involved in planning the organizational objectives. In a team, members feel a sense of ownership for their jobs and unit, because they are committed to values-based common goals that they previously established.
3. **Creativity and Contribution.** In a group, members are told what to do rather than being asked what the best approach would be. Suggestions and creativity are not encouraged. In a team, members contribute to the organization's success by applying their skills, knowledge and creativity to team objectives setting.
4. **Trust.** In a group, members distrust the motives of colleagues because they do not fully understand the role of other members. Expressions of opinion or disagreement are considered troublesome or non-supportive. In a team, members work in a climate of trust and are encouraged to openly express ideas, opinions, disagreements and feelings.
5. **Common Understandings.** In a group, members are so cautious about what they say, that real understanding is not possible. Game playing may occur and communication traps be set to catch the innocent. In a team, members practice open and honest communication. They make an effort to understand each other's point of view.
6. **Personal Development.** In a group, members receive good training but are somehow limited in applying it to the job either by the manager or other group members. In a team, members are encouraged to continually develop skills and apply what they learn on the job.
7. **Conflict Resolution.** In a group, if members find themselves in conflict situations they do not know how to deal with it. Their supervisor/leader may postpone intervention until serious damage is done, i.e. a crisis situation. In a team, members realize that conflict is a normal aspect of human

interaction but they view such situations as an opportunity for new ideas and creativity. They try to manage conflict quickly and constructively.

8. **Participative Decision Making.** In a group, members may or may not participate in decisions affecting the team. Conformity often appears more important than positive results. In a team, members participate in decisions affecting the team but understand their leader must make a final verdict whenever the team cannot decide, or an emergency exists. Positive win/win results are the goal at all times.
9. **Clear Leadership.** In a group, members tend to work in an unstructured environment with undetermined standards of performance. In a team, members work in a structured environment, they know what boundaries exist and who has final authority. The leader sets agreed high standards of performance and he/she is respected via active, willing participation.
10. **Commitment.** In a group, members are uncommitted towards excellence and personal pride. Performance levels tend to be average. Staff turnover is high because talented individuals quickly recognize that:
    (a) personal expectations are not being fulfilled
    (b) they are not learning and growing from others and
    (c) they are not working with the best people.
    In a team, only those committed to excellence are hired. Everyone works together in a harmonious environment.

Why people join groups or work teams

People naturally tend to join groups, being usually surrounded by others with comparable values, this fact reinforce their own value system.

Another reason to join groups might be that groups give people an informal status, which can be the feeling of belonging to a distinct unit. Group membership also provides a level of individual security; members feel that they are equal with the others. Individuals are better listened within a group. Recognition, participation and communication needs are higher satisfied in a group.

To conclude, specific needs of group members are better met in a group than in the whole organization.

The literature describes several types of groups according to a set of criteria.

Formal groups (work team) is created by an organization in order to achieve a certain goal, being recognized and receiving full support from the organization. Under this category, several forms of groups may exist:

- Functional group – consisting of a manager and all its subordinates. In a formal organization, each department may be considered a functional group.
- Operational or task-dedicated group – specially created to achieve a clear task, has a temporary basis and the group is suspended once the task is finished.
- Permanent Committee – it is actually an operational committee having the responsibility to solve problems that appear periodically in a certain field of activity. A committee is not specific only to an organization; there are also inter-organizational committees (for instance inter-ministerial committees).
- Consultative Group – a temporary group having as main task to make recommendations for a certain topic. Usually includes persons from different sectors.
- Self-managed group – it represents the group that has to achieve a certain task without being regularly monitored. Its members have full responsibility, being governed by a spirit of solidarity. They make their own decision on the distribution of tasks inside the group, they establish the working hours, and they are doing the performance self-evaluation.
- Project teams – nowadays, projects often require that people work together in order to accomplish the project goal. Members of these teams might belong to different groups, but receive clear assignment to activities for the same project; thereby outsiders may view them as a single unit.

Informal groups arise from natural attractions among individuals for social reinforcement or other benefits. They seldom share the organizational objectives and have a temporary basis. This type of group is created by people and not by organizations. Informal groups are not formally recognized by the organization and do not represent a unit in the organizational chart.

There are several types of informal groups:

- Group of interests – created usually to facilitate the achievement of group goals;
- Group of friends – created most frequently to satisfy the social needs of its members;
- Group of support – members are supporting each other to meet common needs;
- Virtual group – it's the most modern type, facilitated by the computer and INTERNET connection. This is a group of people who work interdependently and with shared purpose across space, time, and organization boundaries using technology to communicate and collaborate. Virtual

team members can be located across a country or across the world, rarely meet face-to-face, and include members from different cultures (6).

When a group in an organizational context embarks upon a process of self-assessment in order to estimate its own effectiveness and thereby improve performance, it can be argued that it is engaging in team building process.

The process of team building includes:

- clarifying the goal, and building ownership across the team and
- identifying the inhibitors to teamwork and removing or overcoming them.

Self-assessment means that a team is trying to find out both:

- its current strengths as a team
- its current weaknesses

Roles within a team

Belbin's book Management Teams (7) presents the conclusions from his work, studying how the members of teams interacted during business games run at Henley Management College. One of his key conclusions was that an effective team should have members that cover nine key roles in managing the team.

Based on Belbin's model of nine team roles, managers of organizations that are building working teams would be advised to ensure that each of the roles can be performed by a team member. Some roles are compatible and can be more easily carried out by the same person; some are less compatible and are more likely to be done well by people with different skills.

The roles actually represent tasks and functions in the self-management of the team's activities. Belbin developed a test to identify individual team roles. This test is presented as Exercise 1, in this way the students can evaluate their own skills in fulfilling Belbin's roles.

The nine roles described by Belbin are:

Co-coordinator

The Co-coordinator ensures that all members of the team are able to contribute to discussions and decisions of the team. Their concern is for fairness and equity among team members. Those who want to make decisions quickly, or unilaterally, may feel frustrated by their insistence on consulting with all members, but this can often improve the quality of decisions made by the team.

Shaper

The shaper is full of drive to make things happen and get things going; a dynamic team-member who loves a challenge and thrives on pressure. In doing this, they are quite happy to push their own views forward, do not mind being challenged and are always ready to challenge others. The shaper looks for the pattern in discussions and tries to pull things together into something feasible, which the team can then get to work on. This member possesses the drive and courage required to overcome obstacles.

Plant

This member is the one who is most likely to come out with original ideas and challenge the traditional way of thinking about things. Sometimes they become so imaginative and creative that the team cannot see the relevance of what they are saying. However, without the plant to scatter the seeds of new ideas the team will often find it difficult to make any progress. The plant's strength resides in providing major new insights and ideas for changes in direction and not in contributing to the detail of what needs to be done. Although they sometimes situate themselves far from the other team members, they always come back to present their 'brilliant' idea.

Resource investigator

The resource investigator is the group member with the strongest contacts and networks, and is excellent at bringing in information and support from the outside. Whatever the team needs, the Resource Investigator is likely to have someone in their address book that can either provide it or know someone else who can provide it. This member can be very enthusiastic in pursuit of the team's goals, but cannot always sustain this enthusiasm. Being highly driven to make connections with people, the Resource Investigator may appear to be flighty and inconstant, but their ability to call on their connections is highly useful to the team.

Implementer

The individual who is a company worker is well organized and effective at turning big ideas into manageable tasks and plans that can be achieved. The Implementer is the practical thinker who can create systems and processes that will produce what the team wants. Such individuals are both logical and disciplined in their approach. They are hardworking and methodical but may have some difficulty in being flexible. Being strongly rooted in the real world, they may frustrate other team

members by their perceived lack of enthusiasm for inspiring visions and radical thinking, but their ability to turn those radical ideas into workable solutions is important.

Team worker

The team worker is the one who is most aware of the others in the team, their needs and their concerns. He is concerned to ensure that interpersonal relationships within the team are maintained. They are sensitive and supportive of other people's efforts, and try to promote harmony and reduce conflict. They may be the first to approach another team member who feels slighted, excluded or otherwise attacked but has not expressed their discomfort. Team workers are particularly important when the team is experiencing a stressful or difficult period. The Team Worker's concern with people factors can frustrate those who are keen to move quickly, but their skills ensure long-term cohesion within the team.

Completer Finisher

As the name suggests, the completer finisher is the one who drives the deadlines and makes sure they are achieved. They usually communicate a sense of urgency, which push other team members into action. They are conscientious and effective at checking the details. Completer finisher has a great eye for spotting flaws and gaps and for knowing exactly where the team is in relation to its schedule. Team members who have less preference for detail work may be frustrated by their analytical and meticulous approach, but the work of the Completer Finisher ensures the quality and timeliness of the output of the team.

Monitor evaluator

A sober, strategic and discerning member, who tries to see all options and judge accurately. They have a strategic perspective and can judge situations accurately. The monitor evaluator can be overcritical and is not usually good at inspiring and encouraging others. However, this member contributes a measured and dispassionate analysis and, through objectivity, stops the team committing itself to a misguided task.

Specialist

This person provides specialist skills and knowledge and has a dedicated and single-minded approach. They can adopt a very narrow perspective and sometimes fail to see the whole picture.

Group development process

Many teams go through a life-cycle of stages, firstly identified by Bruce Tuckman in 1965 as: forming, storming, norming and performing. This model has become the basis for further models of group development, principally adding the 5th phase of adjourning. The team life-cycle is summarized in Figure 1.

Forming

The process of team formation calls for the individuals to come together. During this phase, group members are exploring new relationships and receiving new responsibilities. The team meets and learns about the opportunity and challenges, and then agrees on goals and begins to tackle the tasks. Team members tend to behave quite independently. They may be motivated but are usually relatively uninformed of the issues and objectives of the team. Team members are more likely to be characterized by formality, politeness, silence and tentative interactions. Mature team members begin to model appropriate behavior even at this early phase. Sharing the knowledge of the concept of "Teams - Forming, Storming, Norming, Performing" is extremely helpful to the team. Supervisors of the team tend to need to be directive during this phase.

The forming stage is important because in this stage the members get to know each other and make new friends. This is also a good opportunity to see how each member of the team works as an individual and how they respond to pressure.

Storming

Due to the fact that the group members become more and more comfortable with the idea of belonging to a certain group, they start to identify their own place within the group. They want to start influencing the group norms, roles (e.g. leadership) and procedures.

In this stage different ideas compete for consideration. The team addresses issues such as what problems they are really supposed to solve, how they will function independently and together and what leadership model they will accept. Team members open up to each other and confront each other's ideas and perspectives.

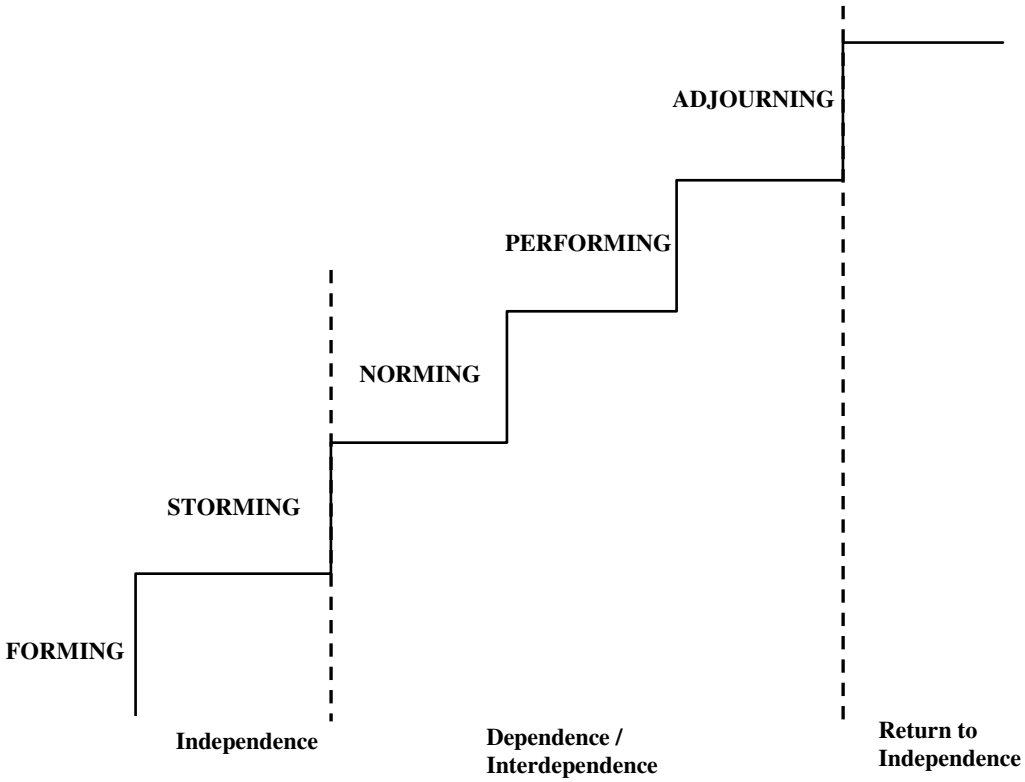
In most of the cases storming can be resolved rather quickly. However, there are cases when the team never leaves this stage. The wisdom of some team members usually determines the end of this stage. Immature team members will begin the "show" to demonstrate how much they know and convince others that their ideas are the best.

The storming stage is necessary to the growth of the team. It can be controversial, unpleasant and even painful to members of the team who are against the conflict. Without tolerance and patience the team

will fail. This phase can become destructive to the team and will lower motivation if allowed to get out of control.

During this phase supervisors of the team may be more accessible, but tend to still need to be directive in their guidance of decision-making and professional behavior.

Figure 1. Group Formation Model



Source: Stewart G., Manz C., Sims H., Team work and dynamics.

Norming

The group becomes more cohesive during this phase because they begin to recognize and respect each other contribution. Team members start to adjust their behavior to the others as they develop work habits that make teamwork seem more natural. Team members often work by agreeing on rules, values, professional behavior, shared methods and working tools. During this phase, team members begin to trust each other. Motivation increases as the team gets more acquainted with the project.

Two categories of norms develop within the team: behavioral and performance. Behavioral ones establish how a person is expected to act and relate to the others. These norms are developed by the group in order to allow conflict resolution, to determine expectations for group-think or to allow divergent thinking and to identify conformity expectations (e.g. be on time, talk in turn etc.). Performance norms are those expectations that are requested by the group from each individual and may include expected contributions to the work. Both types of norms are set to facilitate group process, problem solving and decision making.

Teams in this phase may lose their creativity if the norming behaviors become too strong. Supervisors of the team during this phase tend to be participative more than in the earlier stages. The team members can be expected to take more responsibility for making decisions and for their professional behavior.

Performing

Once the rules were established and recognized, the group can concentrate on the tasks. The high-performing teams are able to function as a unit as they find ways to get the job done smoothly and effectively without inappropriate conflict or the need for external supervision. Team members have become interdependent. By this time they are motivated and knowledgeable. The team members are now competent, autonomous and able to handle the decision-making process without supervision. The focus of the group should be now on results rather than on process, even if work plan revisions might be necessary.

Supervisors of the team during this phase are almost always participative. The team will make most of the necessary decisions. Nevertheless, a change in leadership may cause the team to revert to storming as the new people challenge the existing norms and dynamics of the team.

Adjourning

Once the task(s) for which the group was formed is accomplished, the group can adjourn. It is important to arrange during this phase a formal recognition of the task completed. This might motivate people to embark other tasks and/or groups.

Advantages versus disadvantages of the teamworking

If there are many advantages of the teamworking, there are also disadvantages that should be considered by a manager when building a team.

Generally, a team can better find the best solutions for a complex task than an individual or even a small group of people. Grouping the individual judgments increases the chance of success in solving a problem due to a higher degree of objectivity and to elimination of individual errors.

A definite benefit of working in teams is the positive synergy, meaning that the results achieved by the team are better than the sum of individual results.

However, there are also obstacles that might be encountered during the process of teamworking. Some of the most common are: different organizational practice (when members are coming from different organizations); socio-economic inequalities between the members (age, gender, statute), team objective unclearly stated, etc.

The literature is describing five main dysfunctions of a team (9):

1. Absence of Trust

This dysfunction stems from teams unwillingness to be vulnerable. Team members who are not genuinely open with one another about their mistakes and weaknesses make it impossible to build a foundation for trust.

2. Fear of Conflict

This failure to build trust is dangerous because it sets a brick for the second dysfunction: fear of conflict. Teams that lack trust are not able to engage in vivid debate of ideas. Instead, they resort to roundabout discussions and guarded comments.

3. Lack of Commitment

A lack of healthy conflict might be a problem because it induces the third dysfunction of a team: lack of commitment. Without having expressed their opinions in the course of passionate and open debate, team members rarely, if ever, commit to decisions, though they may simulate agreement during meetings.

Avoidance of Accountability

Because of this lack of real commitment, team members develop an avoidance of accountability, the fourth dysfunction. Without committing to a clear plan of action, even the most focused and driven people often hesitate to call their peers on actions and behaviors that seem counterproductive to the good of the team.

Inattention to Results

Failure to hold one another accountable creates an environment where the fifth dysfunction can thrive. Inattention to results occurs when team members put their individual needs (such as ego, career development, or recognition) or even the needs of their divisions above the collective goals of the team.

There are several factors that can influence the effectiveness of a team. Among them, some are worthy to be mentioned:

- organizational environment – may shape the way of interacting both between team members themselves and between the team and the rest of the organization
- the nature of the task – may influence the way of action, as well as the focus of team efforts. The more complex the task is, the interactions with external environment are more intense and focused.
- team size – there is not a magic number for a team to be effective. It was noticed that increasing the number of members may lead in the first stage to a higher performance. Nevertheless, a continuous increasing of the number of members may lead to the limitation of the team performance or even to a decreased performance.

Examples of successful teams

Sports offer some of the best examples of teamwork. For example a football running back and quarterback's ability are totally dependent on the strength of their offensive line. A basketball player's ability in scoring is mainly dependent on his team's willingness to pass.

Students that succeed in group efforts understand that they must do team projects rather than group projects. There are subtle but very important differences between group and team projects. A team project is when members of the teamwork work interdependently towards the same goal. It is also a team project, when every member in the group feels a sense of ownership of their role. In a group project, members work

independently and are often not working towards the same goal. The members in the group also focus a lot on themselves because they are not involved in the planning of their goals.

Teamwork has also become increasingly acknowledged as an essential skill for employees in companies either small or large. Nowadays increasingly global economy places a premium on teamwork in the work setting. Teamwork has become so valued that many large corporations have developed specific tests to measure potential employees' teamwork abilities. Many companies are even acknowledging this in their job titles by changing the designation of supervisors or managers to "team leader."

Exercises

Task 1

Distribute Belbin questionnaire together with the instructions sheet to the students for the self-assessment of the roles they may have in a team.

Instruction sheet. How to fill-in the questionnaire

1) The purpose of this questionnaire is to assist you in analyzing the role you may play when you work in a team. Be honest with yourself! Without spending too much time on the task, mark the answers to that situation which are true for you at present.
2) Seven situations appear below: sections A-G. After each of them, eight answers are given. You are asked to tick the answer(s) that better describe your behavior under the circumstances provided at the beginning of the section. You can tick maximum 3 answers in each section (you can tick only one if this one is the only one describing your behavior). Repeat this step for each of the seven sections A-G.
3) Come back to section A and evaluate the relative importance (weight) of each of the ticked answers. You should distribute 10 points (no more than 10!) for the answers that you ticked in a section.
   For example, if you have ticked answers 1, 7 and 8 in section A, but you consider that answer 1 is your predominant behavior you can give 5 points for answer 1, another 2 points for answer 7 and finally 3 points for answer 8 (in total: 10 points). If you have ticked only one answer in a section, it automatically receives 10 points.
4) Allocate 10 points for each of the seven sections A-G.
5) When you finished step 4, go to the page containing a summary table.
6) Transfer all the points you allocated under each section in the summary table. Identify your points in the summary table, item by item. The first row is dedicated for section A: considering the example above, you should write 5 in the cell 1___, then 2 in the cell 7___ and 3 in the cell 8____. In this step don't pay attention to the abbreviations from each column. Each line should have a total of 10.
7) After completing step 6, make the total for each of the columns. The highest score in the columns represents the role that better describes your behavior in a team.

Belbin questionnaire

SECTION A

When I am involved in a group project:

1. _____ I can be relied upon to see that all essential work is organized.
2. _____ My general vigilance prevents careless mistakes and omissions being made.
3. _____ I am ready to press for action to make sure that we do not waste time or lose sight of the main objective.
4. _____ I can be counted on to contribute something original.
5. _____ I am able to objectively analyze other ideas, both good and bad ones.
6. _____ I am keen to look for the latest in new ideas and development.
7. _____ I have an aptitude for organizing people.
8. _____ I am always ready to back a good suggestion in the common interest.

SECTION B

I gain satisfaction in a group task because:

1. _____ I like to have a strong influence on decisions.
2. _____ I feel in my element where I can give a task my full attention.
3. _____ I like to feel I am fostering a good working relationship.
4. _____ I enjoy analyzing situations and weighing up all the possible choices.
5. _____ I like to find a field that stretches my imagination.
6. _____ I can get people to agree on a necessary course of action.
7. _____ I am interested in finding practical solutions to problems.
8. _____ I can meet people who may have something new to offer.

9.

SECTION C

When the team has to solve a complex problem:

1. _____ I usually keep an eye on the fields where difficulties might occur.
2. _____ Producing ideas with wider applications is one of my natural assets.
3. _____ I enjoy analyzing situations and weighing up all the possible choices.
4. _____ I am able to coordinate and make effective use of people's skills and capacities.
5. _____ I am always supporting a systematic approach despite possible pressures.
6. _____ I am able to contribute with a new approach for a long-term problem.
7. _____ I am not reluctant to challenge the views of others or to hold a minority view myself.
8. _____ I am always ready to help.

SECTION D

My characteristic approach to daily work in groups is that:

1. _____ I usually want to detect my unclear tasks and objectives
2. _____ I am always ready to express my point of view during the meetings.
3. _____ My ability rests in being able to work with different people whenever I detect they have something of value to contribute to the group.
4. _____ I am keen to detect interesting ideas/people.
5. _____ I can usually find a line of argument to refute unsound propositions.
6. _____ I am able to find possible associations between elements where others cannot detect them.
7. _____ Being very busy produces me a real satisfaction.
8. _____ I have a strong interest in getting to know colleagues better.

SECTION E

If I am suddenly faced with a difficult task with limited time and unfamiliar people:

1. _____ My imagination is often frustrated due to my team work.
2. _____ I believe I have the abilities to reach the consensus.
3. _____ My feelings rarely interfere with my judgment.
4. _____ I am striving to build an effective structure.
5. _____ I am able to work with very different people, despite their personal skills or look.
6. _____ I am ready to face temporary unpopularity if it leads to worthwhile results in the end.
7. _____ I usually have good professional contacts.
8. _____ I feel that I have a natural sense of urgency.

SECTION F

When a sudden new project appears:

1. _____ I start searching possible ideas and perspectives.
2. _____ I am anxious to finalize in the best way my current tasks before starting the new project.
3. _____ I start studying the new project in a careful and analytical way.
4. _____ I am ready to take the lead in involving other people if necessary.
5. _____ I have an independent and innovative position related to most of the possible situations.
6. _____ I would be prepared to take a positive lead if I felt the group was making no progress.
7. _____ I have a positive reaction to all the initiatives of my colleagues.
8. _____ e) I find it difficult to get started unless goals are clear.

SECTION G

What I believe I can contribute to a work team or group:

1. _____ I have the capacity to design a good action plan to achieve a complex task.
2. _____ I might be perceived too analytical, but I usually get very close to achieve the task.
3. _____ A wide network of contacts is important for my work.
4. _____ I am apt to go into details.
5. _____ I am trying to influence the group meetings.
6. _____ I have a clear vision on good ideas and tools that might help the work.
7. _____ I believe my capacity for judgment can help to bring about the commonly agreed decisions.
8. _____ I have good relations with everybody and work hard for the team.

Summary scoring table

	SH	CO	PL	RI	ME	IMP	TW	CF
A	3____	7____	4____	6____	5____	1____	8____	2____
B	1____	6____	5____	8____	4____	7____	3____	2____
C	7____	4____	6____	2____	3____	5____	8____	1____
D	2____	3____	6____	4____	5____	1____	8____	7____
E	6____	5____	1____	7____	3____	4____	2____	8____
F	6____	4____	5____	1____	3____	8____	7____	2____
G	5____	7____	6____	3____	2____	1____	8____	4____
TOTAL								

In the end discuss the results. Advice the students to repeat Belbin test after 1 year when they can be involved in another group or project. Results may be different.

Task 2

Group Survival Scenario Exercise

"Lost at Sea"

With your private yacht slowly sinking after a fire of unknown origin, you are adrift in the South Pacific, "Lost at Sea," approximately 1000 miles south-southwest from the nearest land. You have a serviceable rubber life raft with oars large enough for yourself and crew. You and crew together have 1 package of cigarettes, several books of matches and 5 one dollar bills. You all also have 15 additional items. The exercise problem to be solved is to rank these 15 additional items by considering their survival value.

Survival items to be ranked:

1. a sextant
2. a shaving mirror
3. a quantity of mosquito netting
4. a 5 gallon can of water
5. a case of army rations
6. maps of the Pacific Ocean
7. a floating seat cushion
8. a 2 gallon can of oil/petrol mixture
9. a small transistor radio
10. 20 square feet of Opaque plastic sheeting
11. shark repellent
12. one quart of 160 per cent proof rum
13. 15ft nylon rope
14. 2 boxes of chocolate bars
15. a fishing kit

Student's tasks

The class of students should be divided into groups of 4 and then given 10 minutes to individually score the items.

After this time, the team members should discuss and decide on the team's priority list. A maximum of 20 minutes should be allocated for the section.

When the second phase is complete a pre-prepared answer sheet should be revealed and they should score their individual and teams answers.

A discussion should then take place where individuals should compare their individual score against the team looking at why the scores are different; what changed their minds, how where they influenced etc.

Then, the score of experts should be exposed and see if individual scores were closer to the experts opinion or the team scores were better.

The list of priority items given by the experts was the following:

According to the experts (US Coastguard), the basic supplies needed when a person is stranded mid-ocean are articles to attract attention and articles to aid survival until rescue arrives. Articles for navigation are of little importance since even if a small life raft were capable of reaching land, it would be impossible to store enough food and water to survive for the requisite amount of time.

Without signaling devices, there is almost no chance of being spotted and ultimately rescued. Furthermore, most rescues occur within the first 36 hours and a person can survive with only a minimum of food and water during that period.

So, the following is the order of ranking the items in their importance to your survival:

1. Shaving Mirror	Critical for signaling
2. 2 gallon can of oil/petrol mixture	Critical for signaling. The mixture will float on water and could be ignited with one of

	the £5 notes and a match.
3. 5 gallon can of water	Necessary to replenish fluids lost through perspiration.
4. One case of army rations	Basic food intake
5. 20 square feet of opaque plastic	Can be utilized to collect rain water and provide shelter from the elements
6. 2 boxes of chocolate bars	Reserve food supply
7. Fishing kit	Ranked lower than the chocolate as 'a bird in the hand is worth two in the bush'. There is no guarantee you will catch any fish.
8. 15ft of nylon rope	Could be used to tie people or equipment together to prevent it being washed overboard.
9. Floating seat cushion	A life preserver if someone fell overboard.
10. Shark repellent	Enough said.
11. One quart of 160 per cent proof rum	Contains 80% alcohol, which is enough to be used as an antiseptic for any injuries, otherwise of little value – would cause dehydration if ingested
12. Small transistor radio	Of no use without a transmitter. You would also be out of range of any radio station.
13. Maps of the Pacific Ocean	Worthless without navigation equipment. It does not matter where you are but **where the rescuers are!**
14. Mosquito netting	There are **NO** mosquitoes in the midpacific ocean.
15. Sextant	Useless without the relevant tables and a chronometer.

Pre-prepared answer sheets:

Objects	**Stage I Individual order**	**Stage II Team order**	**Stage III Experts' order**	**Stage IV Difference between I-III**	**Stage V Difference between II-III**
Sextant					
shaving mirror					
mosquito netting					
5 gallon can of water					
army rations					
maps of the Pacific Ocean					
floating seat cushion					
2 gallon can of oil/petrol mixture					
transistor radio					
20 square feet of Opaque plastic sheeting					
shark repellent					
one quart of 160 per cent proof rum					
15ft nylon rope					
2 boxes of chocolate bars					
fishing kit					
				Sum of column IV scores (do not take into account +/- signs)	Sum of column V scores (do not take into account +/- signs)

To evaluate your team performance, fill-in the last sheet:

Stage VI. Individual mean score (divide the Sum of column IV scores with 15)
Stage VII. Team mean score (divide the Sum of column V scores with 15)
Calculate the GAIN – compare VII and VI. If VII is smaller than VI, then the gain of the team is positive. If VII is higher than VI, then the team gain is negative, your individual thinking was better.

References

1. Longest B.B., Rakich J.S., Darr K., Managing Health Services Organizations and Systems, Health Professions Press, inc. Baltimore, 2000.
2. Mills T.M, The Sociology of Small Groups, 2nd edition, 2. Englewood Cliffs, NJ: Prentice-Hall, 1983.
3. AUPHA (Association of University Programs in Health Administration) Users guide, 1996.
4. Katzenbach J.R., Smith D.K.. The Wisdom of Teams: Creating the High-performance Organization. Boston: Harvard Business School, 1993.
5. Diamond Linda Eve, Teambuilding That Gets Results: Essential Plans and Activities for Creating Effective Teams (Quick Start Your Business) (Paperback), January 2007.
6. Kimble C., Barlow A, Feng L, Effective Virtual Teams through Communities of Practice, Department of Management Science Research Paper Series, 00/9, University of Strathclyde, Strathclyde, UK, 2000.
7. Belbin M., Management Teams. London; Heinemann, 1981.
8. Stewart G., Manz C., Sims H., Team work and dynamics. New York: John Wiley&Sons Inc., 1999.
9. Lencioni P., The Five Dysfunctions of a Team: A Leadership Fable, Jossey-Bass Publisher, 2002.

HEALTH INVESTIGATION: ANALYSIS – PLANNING - EVALUATION A Handbook for Teachers, Researchers and Health Professionals	
Title	**Project management**
Module: 2.43	ECTS (suggested): 1.0
Author(s), degrees, institution(s)	Silvia-Gabriela Scintee, MD, MSc, PhD General Deputy Director, National School of Public Health, Management and Professional Development Bucharest, Romania Adriana Galan, Public Health Consultant National Institute of Public Health, Bucharest, Romania
Address for Correspondence	Silvia-Gabriela Scintee National School of Public Health, Management and Professional Development 31 Vaselor Str, 021253 Bucharest Romania Tel. +40212527893 E-mail: gscintee@ispb.ro; sgscintee@snspms.ro
Keywords	Project development, project implementation, project management, project proposal writing.
Learning objectives	After completing this module, students should be able to: • identify what it is needed to know and do to set up a project; • identify the needs and set up the priorities; • prepare a Preliminary Brief; • plan and schedule a project; • set up a monitoring and evaluation plan; • develop a project proposal.
Abstract	This module concerns Basic concepts of Project Management, Initiation Phase, Preliminary brief, Detailed Planning, Scheduling, Implementation and Completion. At the end of this module, the students will design a draft of a project proposal.
Teaching methods	Lectures, interactive presentation of key concepts (overheads or PowerPoint presentation), group discussions, groups' assignments. Work in small groups (4-5 persons) and an overhead will be presented by each group after each assignment.
Specific recommendations for teachers	It is recommended that this module is organized within 1.0 ECTS credit. The work under supervision is consisting from lecture (5 hours), supervised assignment solving (5 hours), while individual work is related to collect data and to prepare the project proposal draft (20 hours).
Assessment of Students	Each student will present a draft of a project proposal, using a ppt presentation and will answer questions related to project management.

PROJECT MANAGEMENT

Silvia-Gabriela Scintee, Adriana Galan

Project is not a new concept. Projects have been carried on since the inception of the organized human existence and less complex projects are very common in our daily life. Any work which has a beginning and an end, is planned and controlled and creates change can be called "project" (1).

Very often "project" is considered synonymous with "programme". Still, there is a difference between the two terms. A programme is more exhaustive than a project and has larger time limits. A programme can have more projects as component parts.

Projects are classified under four main headings (2):

1. Industrial projects (civil engineering, construction, petrochemical, mining and quarrying projects) - usually large projects, requiring massive capital investment and rigorous management, that incur special risks as the implementation phase is conducted remote from project manager's office;
2. Manufacturing projects (production of equipment or machinery) - these are conducted in a factory, but sometimes requiring work away from the company for installation, customer training, subsequent service and maintenance;
3. Management projects - arise in each organization as part of its work or when a change is envisaged; examples of such projects: restructuring the organization, relocating the headquarter, refurbishing an office, planning a training session or conference, introducing new service, introducing a new computer system;
4. Research projects - unlike other types of projects their final objectives are difficult or impossible to define.

The characteristics of the project

Regardless its type, the project has five common characteristics (3): it creates change, it has various goals and objectives, it is unique, it is limited in time and scope and it involves a variety of resources.

1. A project creates change

When a project is conducted the routines and regular work within an organization is disrupted by unfamiliar, new activities. This could lead to resistance from the staff as people do not like to have their existing work altered. More than that, those working in the project have to report to the supervisor of the routine work and also to the supervisor for the project work. Other sources for resistance could be: conflict of interests, low tolerance to change, different perception of the need to change, misunderstanding and lack of trust.

The Project Manager could use various methods of change management to face this resistance in accordance with how much time, money and power he has. Examples of such methods are:

- education and communication – the best method which unfortunately takes a lot of time and money;
- involvement of people in project development – also good but still takes time;
- supporting people to facilitate change – takes also time;
- negotiating with people – takes less time, but a compromise has to be reached;
- manipulating people – quick and cheap method, but can fail if people feel like being used;
- coercion – the quickest and the cheapest, but in the same time has the highest probability to fail.

2. A project has various goals and objectives

There are three types of goals and objectives for any project (3):

- performance and quality – the end result of the project must fit the purpose for which it was intended;
- budget – the money spent on the project must correspond to the authorised expenditure;
- time of completion – all stages of the project should take place at their specified dates and total completion date should correspond to the planned finish date.

The Project Manager should find a balance between these three attributes: time, quality and cost. If the project finishes before the planned completion date, money might be lost. If the project is extended beyond its scheduled finish date is likely to have increased costs. In both cases the quality might suffer.

3. A project is unique

There are not two identical projects. Despite the existence of a standard methodology for project development and of a same basic procedure no matter the complexity of the project, the work content of every project varies.

The Project Manager needs to develop a plan taking into account the particular circumstances that is both strong and flexible enough to accommodate changes in those circumstances.

4. A project is limited in time and scope

These are the main characteristics that make the difference between the project and the programme. A programme is not necessarily limited in time and its scope is more comprehensive. A project is limited in time and scope, having a beginning and an end very well defined.

A great deal of the Project Manager's effort is focused on the completion of the project at the scheduled finish date. There are a lot of tools that can be used in time planning, from timetables – the simplest, which represent a list of activities with their starting and finishing dates, to Gantt Chart and Critical Path - more complex methods that take into account the dependency degree between activities.

5. A project involves a variety of resources

When a project resource planning is discussed, most people will think of resources first in terms of money. But resources are also: people, equipment, materials and time. It is very difficult to forecast the precise quantity of resources and the moment when these will be used. Still, the necessary resources should be estimated and scheduled. It is also necessary to specify how these resources will be obtained.

As there are a lot of factors that could impede the utilization of resources according to the schedule, the Project Manager must periodically evaluate the progress and, if necessary, re-schedule resources.

The above characteristics of the project have implications for the project management that is defined as "the process by which the project manager plans and controls the tasks within the projects and the resources on which the organisation draws to carry out the projects" (1).

The phases of the project

All projects may be planned and carried out in the following four phases, known also as "the project's life cycle": initiation, planning, implementation, and completion.

Initiation. Projects arise because of a need. So, in the initiation phase there are determined: the need for the project development, the terms of reference (what has to be done, what would be the expected results), the feasibility of the project and also it is created a workable environment for the project. This is considered the most important phase in the whole project (1), even if is the shortest one usually taking no more than 5% of the project lifetime.

Planning. In the planning phase, which usually takes 20% of the project duration, the tasks, resources, effects and needs of the project are examined in depth (1). Planning is under the responsibility of the Project Manager, either done by himself in isolation or by a planning team. During this phase it is decided what should be done, by whom, at what point in time and with what resources in order to reach the project's objectives. It is important in the planning process to forecast the potential constraints that might affect the implementation phase and to design strategies for overcoming them.

Implementation. Implementation is the longest phase of a project (60% of the project duration) in which the project plan is put into operation. The implementation process is monitored and controlled in order to ensure the obtaining of quality results on time and within budget. Monitoring is a continuous oversight of the project execution that assists in its supervision and assures that it proceeds according to plan. On the basis of controlling the project, progress is checked against the plan and corrective action is taken where necessary.

Completion. In the last phase the whole project is reviewed, the final report is presented and the resources are re-allocated. This phase usually takes 15% of the project duration.

1. Initiation (pre-planning) phase

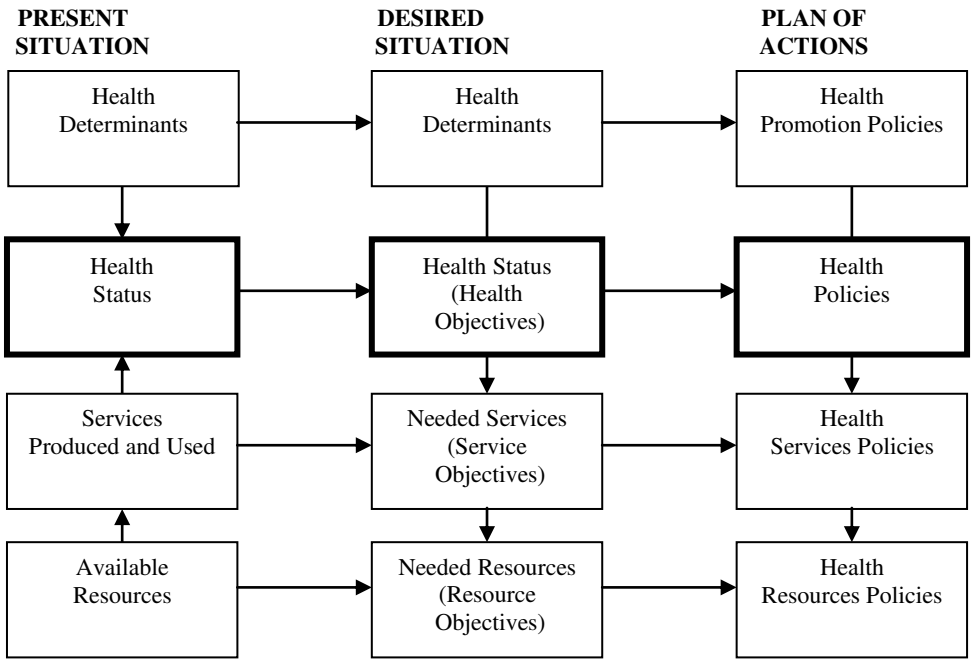
This is generally considered the most important phase in any project. During this phase, **what should be done** under the project is determined (1). This phase includes usually the following steps:

1. Situation analysis
2. Health problems identification
3. Priority setting
4. Establish goal and objectives
5. Feasibility study

1.1. Situation analysis

Situation analysis represents the first step of the pre-planning phase for any project. It could represent an assessment of the health status of the population (can be a "target" population) or an assessment of the health care system in relation with the internal and external environment. According to R. Pineault (4), the general framework of a situation analysis can be conceptualised as presented in the figure 1.

Figure 1. The general framework of a situation analysis



Source: adapted after R. Pineault

The main goal of this step is to define and establish valid criteria for the identification of priority health problems. Another important goal is to provide data and information necessary to design objectives and strategic choices. It also represents a support for the feasibility evaluation.

Data and information collected during this step cover the following domains:

- Assess the internal and external environment (review of economic, social and health objectives and policies);
- Health status and related determinants assessment (mortality and morbidity rates, disability, life expectancy, lifestyle indicators, trends etc);
- Health system assessment (public/private institutions, accessibility for health care, population coverage with services, patient flow within the health care system, etc);
- Resources – human, material and financial.

The main output of this step is represented by a comprehensive document offering a picture of the existing situation.

1.2. Health problems identification

The main goal of this step, involving more or less a subjective judgement, is to obtain a list of health problems. According to R.Pineault (4), a **health problem** represents a deficient health status as perceived by individuals, physicians or communities.

There are several methods described in the literature for problem identification. R.Pineault (4) has described three categories of approaches:

- Based on existing health system indicators
- Based on special surveys
- Based on consensus research

For each approach, he described the methods used in order to identify the health problems. Table 1 presents the methods used within each approach.

Table 1. Problem identification methods

Approach	Methods	Needed Information
Based on existing health system indicators	Socio-demographic (associated to the health status and service utilisation)	Population structure, age pyramid, natality rate, crude mortality rate, fertility rate, average income level, level/rate of poverty, rate of unemployment, level of education
	Health (mortality, morbidity, risk factors and disability)	Crude and specific mortality rates, infant mortality rate, life expectancy at birth and certain ages, standardised mortality rates/ratio Incidence/prevalence rates, hospitalised morbidity Frequency of different risk factors, attributable deaths for certain risk factors, potential years of life lost due to certain risk factors DALY, QALY
	Health services utilisation	Medical visits rate, surgical interventions rate, , number of diagnosis tests(e.g. laboratory, x-rays etc.), number of referrals, hospitalisation rate (number of discharges), average length of stay
	Health resources	Number and types of health care units, population coverage with different types of health care professionals (physicians, nurses, dentists etc), health care expenditures
Based on special surveys	Sampling	Health Interview Surveys (perceived health status) Health Examination Surveys (based on clinical exams)
Based on consensus research	Delphi Technique	Evaluate the opinion of certain experts on prevalent problems in a community. It is based on a group process of judgement, even if the experts don't communicate directly. The experts answer to successive posted questionnaires until sufficient level of consensus is reached
	Nominal Group Technique	It is also a group technique, aiming to identify the problems and to order them according to their importance. Consensus is reached by individual voting on the identified problems.
	Brainwriting Technique	The difference from the Nominal Group Technique is that all the ideas concerning the problems are presented (written on a table) from the very beginning to all participants. It is possible to reach the consensus also by voting or by final discussion.
	Brainstorming Technique	It is mainly useful to generate ideas (mostly recommended for problem analysis and judgement of choices). Experts are invited and encouraged to come up with original ideas.
	Community Forum	Public is invited to express community problems.

Source: adapted after R. Pineault.

In order to judge the identification of one problem, several criteria can be used:

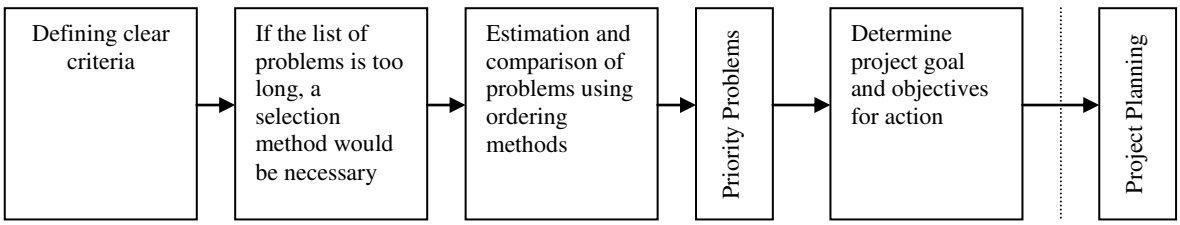
- Problem's dimension (usually its frequency within a population);
- *Problem's severity* (usually measured by premature deaths, potential years of life lost, disability);
- Trends.

1.3. Priority setting

Priority setting means to select those problems identified during the previous step that can be the object of an intervention. It is actually a process of comparisons and decision-making, based on special methods and techniques for ordering the identified problems according to their importance.

The conceptual framework of priority setting process (figure 2) was also described by R.Pineault (4).

Fig. 2 Priority setting conceptual framework



Source: adapted after R. Pineault.

Three main criteria are used in order to prioritise the identified problems:

- *problem's* dimension (incidence/prevalence, premature deaths, avoidable deaths, invalidity, the size of the population at risk, the impact on medical services, family, society, etc);
- intervention capacity (knowledge on the disease/associated risk factors, prevention possibilities);
- existing resources for intervention (existing services, qualified personnel, population accessibility to health services).

There is a wide range of priority setting tools (ranking methods) that can start from a simple grid analysis, and ends with complex methods. Based on a large number of criteria, these tools allow the problems ranking. If the list of identified problems is too long (>40), it would be necessary to shorten this list, using the Selection Method.

Selection method

Its main purpose consists in rejecting the less important problems from the list. The result of selection method is a shorter list of more important problems, and not necessarily a problem ranking.

A selection criteria is established from the beginning. A group of 3-5 experts will select the most important and less important problems during several meetings:

- first meeting: the most important and least important problems are selected from the initial list, and put on separate lists (important and less important problems);
- second meeting: from the remaining list, the first 2 most important problems and the last 2 less important problems are again selected and put on the 2 previous lists;
- third meeting: from the remaining list after the second meeting, the first 4 most important problems and last 4 less important are again selected and put on the 2 lists;
- the process stops when the list of most important problems contains no more than 10 problems.

R.Pineault has grouped the priority setting (ranking) tools into two categories (4):

- specific methods for health planning
- general ranking methods

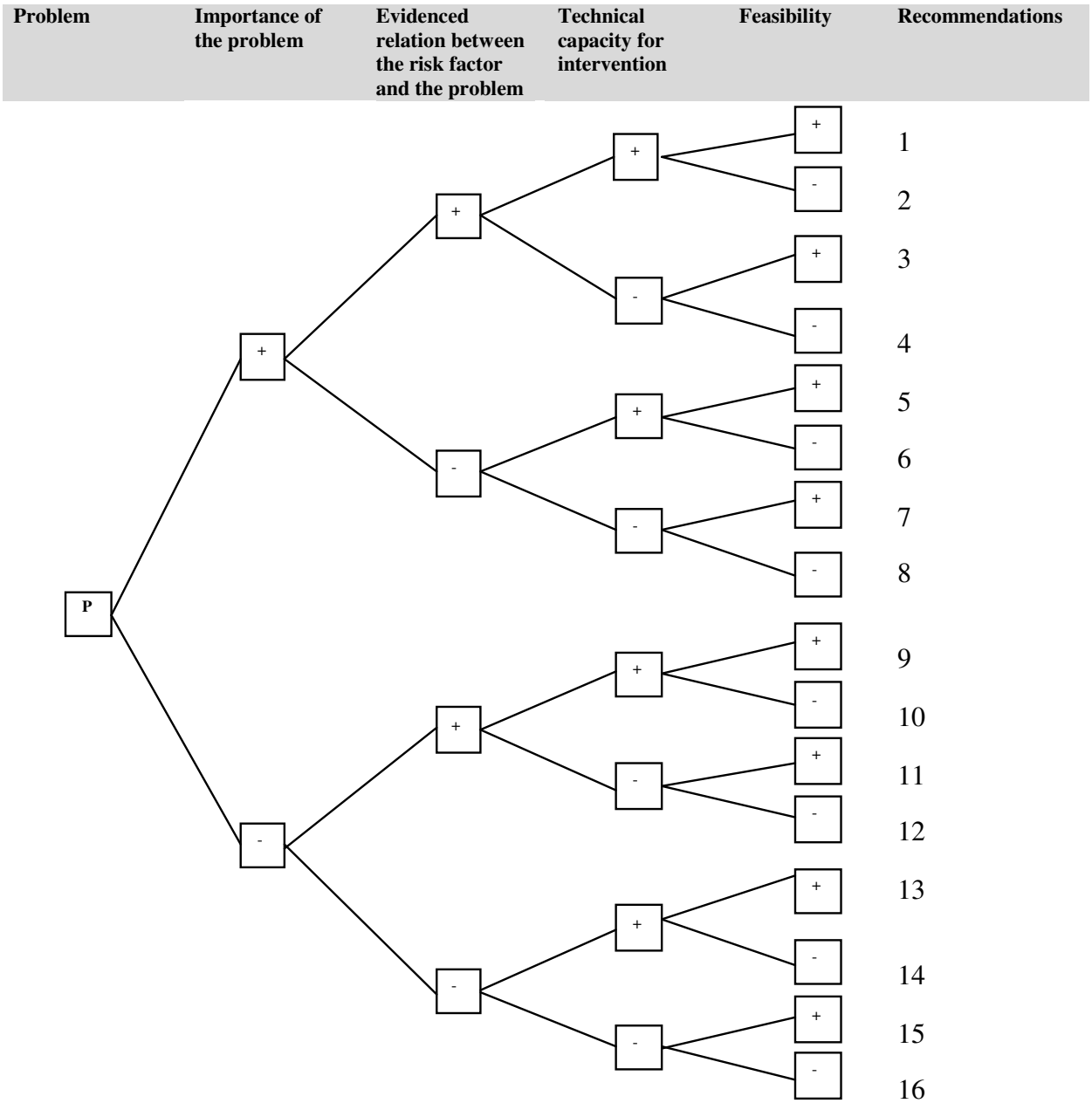
Specific Methods for Health Planning

Within this category, two methods are mostly used: Grid Analysis and Hanlon Method.

Grid Analysis

It allows formulating recommendations on priorities. It takes into account the problem importance, its evidenced relationship with associated risk factors, technical potential for problem solving, and intervention feasibility. The method allows establishing 16 possibilities of recommendations in descending order of priority for each problem. A general Grid Analysis is presented in the figure 3 below.

Fig. 3 Grid Analysis process



Source: adapted after R. Pineault

The results can be summarised as in table 2

Table 2 Summary of a Grid Analysis

	Problem importance	Evidenced relationship with associated risk factors	Technical potential for problem solving	Intervention feasibility	Recommendation from the Grid Analysis
Problem 1	+	+	+	-	2
Problem 2	-	+	+	+	9
Problem n	-	+	+	-	10

Source: adapted after R. Pineault

According to this method, Problem 1 is considered the most priority.

Hanlon Method

It ranks the priorities taking into account 4 components: problem magnitude (A), problem severity (B), solution effectiveness (C) and intervention feasibility (D).

A. Problem magnitude is usually measured by rates or index (a score is assigned for each problem; score values ranges between 1 and 10. Value 10 represents the highest frequency in a population).
B. Problem severity is usually measured by mortality rates, potential years of life lost, DALY, associated costs (a score is assigned for each criteria; score values ranges between 1 and 10. A final score is calculated for each problem, as the average of previous scores. Value 10 represents the most severe situation).
C. Solution effectiveness must measure the availability of resources and technologies able to improve the problem. A score is also assigned for each problem, ranging between 0.5 and 1.5. Value 0.5 indicates that the problem is difficult to be solved, while 1.5 indicates that there are possibilities to solve the problem. It is mostly a subjective judgement.
D. Intervention feasibility is also a subjective judgement taking into account the following components for each problem: pertinence (P), economic feasibility (E), acceptability (A), resources availability (R) and legal framework (L). A score is assigned for each component, 1 means a positive answer, 0 means a negative answer.

A final composite index is computed for each problem based on the following formula:

$$P_{1.....n} = [(A+B) \times C \times D]$$

The highest score corresponds to the most priority problem.

General Ranking Methods

Within this category, the following methods can be mentioned: anchored rating scale, paired comparison and pooled rank.

Anchored rating scale

A linear scale is used, ranging between 0 and 1 (1 = extremely important problem; 0.75 = very important; 0.5 = important; 0.25 = less important; 0 = problem can be neglected). Each expert is asked to place every problem on this scale. Finally, a mean is calculated for each problem, having in the end a hierarchy.

Paired comparison

Problems are compared two by two. During each step, a problem is compared with all the others; for each comparison the most important problem is marked. For each problem it is computed in the end o sum of favourable situations. For example, if there are 5 problems (A, B, C, D, E) to be ranked, the method can be summarised as shown in table 3.

Table 3. Paired comparison ranking

Problem	Paired comparison (selected problem is marked)				Obtained score or percent
A	A B√	A C√	A√ D	A E√	A=1 or 10%
B		B√ C	B√ D	B E√	B=3 or 30%
C			C√ D	C E√	C=2 or 20%
D				D E√	D=0 or 0%
E					E=4 or 40%

Source: adapted after R. Pineault.

Pooled rank

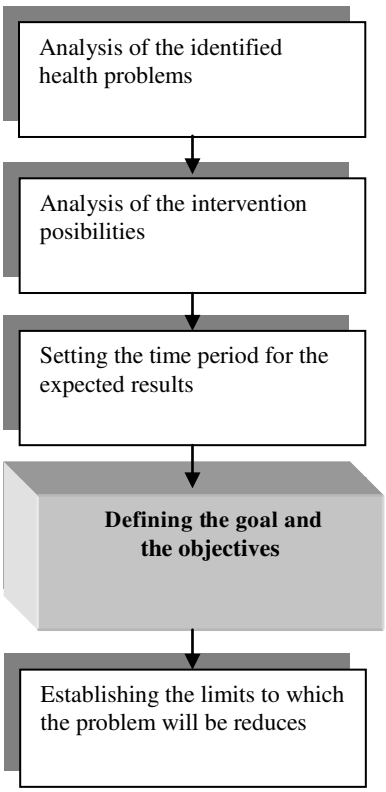
A group of experts is ranking the problems, starting with the most important one (highest rank) and ending with the least important (lowest rank). Each problem receives a rank from each expert. A mean rank is finally computed for each problem.

1.4. Establish goal and objectives

In establishing the goal and the objectives, the following elements should be taken into account: the goal and the objectives of the national health policy, the goal and the objectives of the national health programme adressing the identified problem (if there is one), local health policies, and international health

standards and objectives. There is also necessary to define: the target population, the geographical area, the extent to which the problem can be reduced or solved, and the time during which the problem should be reduced or solved. Stages towards defining the goal and objectives are presented in figure 4.

Fig. 4 Stages towards defining the goal and objectives



A **goal** is a long term result toward a project is aiming. In health, a goal usually refers to the solving or reduction of a health problem. There is not necessary to specify any quantitative outcome or time limits (5).

Ex. "To increase the reproductive health by reducing the number of abortions and undesired pregnancies in students from Bucharest University"

An **objective** is a desired outcome to be reached in a certain period of time. An objective measures the progress towards the stated goal. For this it is necessary to be quantified and to establish time limits. In defining an objective the following have to be specified:

- what will be achieved;
- how much (to what extent);
- when is expected the result;
- who will benefit;
- where is expected the result.

In defining objectives could also be used the acronym SMART (S = specific, M = measurable, A = agreed upon, R = realistic, T = timebound). It is recommended a limited number of objectives (3 – 5). In accordance with to the project complexity there can be established different types of objectives:

– General objective – which would be the result expected at the completion of the project and shows how much the situation will improve;

Ex. "To reduce by 50% the number of abortions and undesired pregnancies in students from Bucharest University, between 2010-*2012"*

– Intermediary objective – measures the progress towards the achievement of the general objective expected at a certain point in time;

Ex. "To reduce by 25% the number of abortions and undesired pregnancies in students from Bucharest University, untill December 31, 2011"

- Specific objectives – represent specific results that would assure the achievement of the intermediary and general objectives;

  Ex. "To increase the information level of the students from Bucharest University in regards with contraceptive methods"
- Operational objectives – that are in fact, the actions to be taken in order to reach the objectives;

  Ex. "To freely distribute 10000 brochures on contraceptive methods to the students from Bucharest University, between January – *June 2010"*

Sequence and interdependence between different types of objectives are shown in figure 5.

1.5. Feasibility study

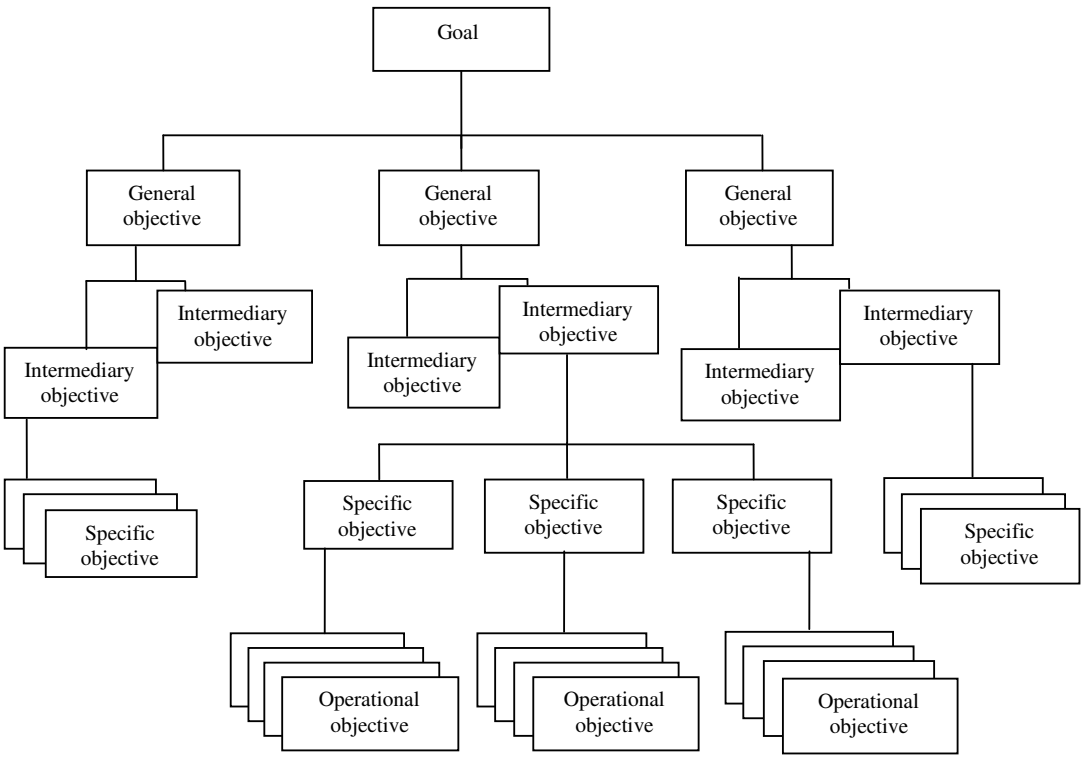
The aim of this step consists in the evaluation of alternative proposed strategies in order to select the best one to be further implemented. The evaluation is focusing on three main aspects (6): political feasibility, technical feasibility, institutional feasibility and financial feasibility.

Political feasibility is focusing on the favourable/unfavourable political environment, on the agreement/disagreement of all key stakeholders involved.

Technical feasibility usually takes into account three aspects:

- provision of requested services needed to achieve the proposed activities (existence and availability of necessary technology);
- the proposed offer of services (meet the population needs?; are the services accessible?; does it attain the target population?);
- impact on health status (do the proposed services improve the health status?; do the services contribute to the achievement of project objectives?).

Fig. 5 Sequence and interdependence between different types of objectives



Institutional feasibility is focusing on:

- estimation of the necessary types of institutions and their geographical distribution, for the achievement of objectives (do they exist?; do they need restructuring / rethinking?; new institutions are needed to be created?);
- staff (existing staff has sufficient skills?; are training sessions necessary?; new staff is necessary to be hired?);
- administrative and managerial capacity (new capacities are needed?; is the logistic support available?).

Financial feasibility takes into account:

- estimation of total costs of necessary resources;
- estimation of running costs of the project;
- identification of possible financial sources.

Preliminary brief

A brief contains the key information about the project, having a multiple use: to proceed a feasibility analysis, to ask for funds, to direct the further planning of the project.

A preliminary brief should include:

- Project name;
- Background (presenting the identified problem and the chosen solution);
- Goal and objectives;
- The expected results;
- The required budget and time;
- Methods of monitoring and evaluation;
- Information about the organization;

2. Detailed planning and scheduling

After objectives setting, a detailed plan of action is developed for each of them. Action plans specify what should be done, by whom, where and when, being the bridge between stated objectives and the practical work. Action plans could be seen as means and methods by which the objectives will be reached.

A project plan should be detailed enough in order to:

- provide a clear image on the activities;
- clarify for the project team the sequence and interdependence of activities;
- facilitate the correct estimation of the necessary resources.

There are described eight steps to be taken for the detailed planning of a project (1):

- identifying the tasks (deciding what has to be done);
- classifying the tasks and placing them in a logical order (some tasks are concerned with running the project, others are concerned with the actual work content of the project);
- studying the implications (how the project could affect the organization policy, what is the impact on the clients, the public, the environment, what is the relationship with other projects);
- estimating resource requirements;
- identifying the project hierarchy;
- clarifying the levels of authority (and setting clear areas of responsibilities for each person);
- setting up the procedures needed to monitor and control the project;
- setting ground rules (informing the team of what is expected as a group norm).

In order to schedule the work content and resources of a project there are a lot of tools that a manager could use. The most known are: the Gantt chart and the Critical Path Method.

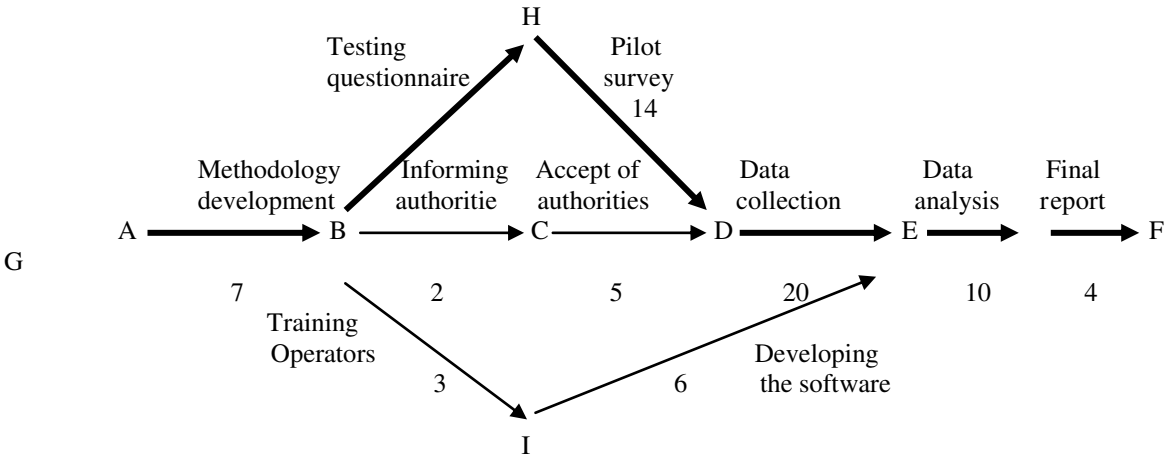
The Gantt chart is recommended for the uncomplicated projects. A Gantt chart (figure 6) is a simple display of tasks (listed in the first column) together with their duration of accomplishment (presented as horizontal bars alongside each task). The time periods could be presented either in days, weeks, months, quarters or years.

For more complex projects in which the dependencies between activities needs to be shown a PERT diagram could be used. PERT (Program Evaluation and Review Technique) is a network tool relating tasks to each other on the basis of time and precedence and producing a critical path through the project (7). Each activity is represented by an arrow, on which the activity is described together with the estimated duration (figure 7). Critical path is the longest path through the network of tasks that defines the duration of the project (7). For this path the Project manager has to worry about as any delay of an activity could lead to the delay of the project end.

The financial resources should also be planned. This is done by using the budget. Budget estimation is very important for a project because:

- is one of the essential elements a funding agency is looking for;
- represent a basis for the financial control that will compare the plan with its execution;
- helps in choosing the most cost-effective projects, attaining the allocative efficiency;
- allows a better resource allocation within a project, attaining the operational efficiency.

Fig. 7 Example of a critical path



A-B-H-D-E-F-G = 7+7+14+20+10+4 = **62 days**
A-B-C-D-E-F-G = 7+2+5+20+10+4 = 48 days
A-B-I-E-F-G= 7+3+6+10+4 = 30 days

In this example the first path has to be taken as critical path.

Tasks	April	May	June	July	August	September
Methodology Development						
Testing the questionnaire validity						
Informing District Health Authorities and training of operators						
Data collection						
Data analysis						
Development of the final report and dissemination of the results						

In order to estimate the budget it is necessary to:

- list all types of required resources for each activity;
- determinate the quantity of each type of resources;
- estimate the unitary cost for each type of resources;
- calculate the total cost of each type of resources;
- discount future costs if the project duration is more than one year.

The costs of each activity are usually presented in four expenditure categories:

- personnel (like salaries, training, per diem, etc);
- equipment and materials (including also maintenance costs);
- facilities (ex. renting, modifying or building a new office);
- support expenditures.

A special category is represented by incidentals which usually should not exceed 10% of the total cost of the project should be justified. The budget should also contain the sources of funding. These could be represented by the organization's own funds or there could be multiple financing organizations. Each source will be specified for in separate columns. The estimated costs could be presented like in the table 4.

Table 4. Example of project breakdown of costs

Expenditure Categories	Activity description	Cost per activity	Own funds	Requested funds
1.Personnel - salaries - accomodation - perdiem - transport	- project coordinator salary 100$/month x 12 months - 15 participants in a training course held in Bucharest x 5 days x 30$per diem - 2 trainers x 5 days x 50$feel/day	1200$ 2250$ 500$		1200$ 2250$ 500$
SUBTOTAL		**3950$**		**3950$**
2. Equipments and materials	- multiplying course materials 5$/participant/day x 5 days x 15 participants	375$		375$
SUBTOTAL		**375$**		**375$**
3. Facilities	- classroom rent 100$/day x 5 days	500$	500$	
SUBTOTAL		**500$**	**500$**	
4.Support expenditures	- communications	300$	300$	
SUBTOTAL		**300$**	**300$**	
5. Incidentals (reimbursed on the basis of receipts)				
TOTAL GENERAL		**5125$**	**800$**	**4325$**

A funding agency might have its own administrative procedures, so before submitting a project the agency should be contacted and should be asked about the necessary documents and the recommended budget format. The plan is many times negotiated with the funding organization. Usually the project should be in accordance with donors' policies and priorities. When deciding to fund a project a financing organization is mainly interested in:

- project justification;
- technical capacity for running the project;
- compatibility with other projects;
- measurable and acceptable benefits;
- detailed and justified costs;
- sustainability (how the impact of the project will be continued after the project funding has ceased);
- a clear monitoring plan;
- previous experience of the applicant;
- collaboration with other partners;
- multiple financing sources.

3. Implementation

The implementation phase consists in putting the project plan into operation once all approvements and authorizations have been received. The plan should be flexible as even after being approved, in the implementation phase, changes might inevitably occur because of the internal or external factors. Examples of internal factors could be: a key person that leaves the team, poor communication on somebody's part, delays in equipment procurement or in funds release. External factors are less under the Project Manager's control. Examples of external factors are: partners who leave the project, change in donor's policy, change in health policy or legislation, change in organization's structure.

Implementation is initiated by the Project Manager and the other authorities responsible for the project by developing the job description for the Project Manager. Than the project team will be completed and the team roles will be assigned after assuring that everyone has a clear vision about the project and, if necessary, after training the team members for working together. The project plan will be reviewed and

detailed as much as possible and tasks and responsibilities will be assigned for each member of the team, as well as the relationships between them. It is very important to set clear responsibilities and communication lines and to establish the authority levels in order to avoid overlaps, misunderstandings or delays in completion of tasks.

Over the implementation, the Project Manager should ensure that the necessary resources will be released on time for each activity. He should forecast the possible risks for not getting the resources in due time and should develop strategies to overcome these problems.

An ongoing process during the implementation is monitoring. Monitoring focuses on periodic measurement of workplan progress and achievement of intermediate project milestones. Properly performed, monitoring provides current supervision and timely opportunities for remedial action (7). Factors to consider in determining the scope and magnitude of the project monitoring are:

- cost of the project;
- previous experience of the implementing team;
- manager's familiarity with and confidence in the implementing team;
- complexity of the project;
- potential for injury to the project due to delays in both reporting and responding.

If monitoring is a method of ongoing review and measurement of the project to gauge its progress relative to its objectives and to plan continual improvements to both activities and management, evaluation takes a broad view of the projects activities, measuring the project's success and effects and showing what difference will the project make (8). F. Champagne (6) has defined the evaluation process as being a judgement on any activity, provided service or project component. The judgement is always based on some criteria and norms (normative evaluation) – mostly used for project evaluation - or on some scientific methods (evaluative research). During project implementation, evaluation can be done as internal and external audit (operational evaluation) which can propose ongoing corrections.

Usually, any evaluation is focusing on the three classical components:

Structure – the resources used by the project are evaluated:

- Human (number, level of competence, existence of incentives);
- Material (quantity and quality);
- Financial (budget);
- Characteristics of the responsible organization: size, type, affiliations, degree of specialisation.

Process – is focusing on the following aspects:

- Project planning (appropriateness and adequacy of activities);
- Project monitoring (existence of periodic and final reports);
- Project organization (leadership, human relationships, responsibilities);
- Project stage related to established deadlines and budget.

Outputs/Outcomes – is focusing on specific results achieved by the project as compared with established objectives:

- Provided activities/services in order to achieve the objectives;
- Obtained indicators;
- Intervention impact (follow-up of an indicator after the end of the intervention).

During the implementation stage reports will be required. Reporting allows Project Managers to share the findings of the project through monitoring and evaluation, requiring periodic documentation of the project progress. A stage report includes financial updates, implementation status report and periodic evaluations.

4. Completion

During this phase, the final project evaluation usually takes place. This is called a-posteriori evaluation and it measures the level of project objectives achievement, project impact on target population.

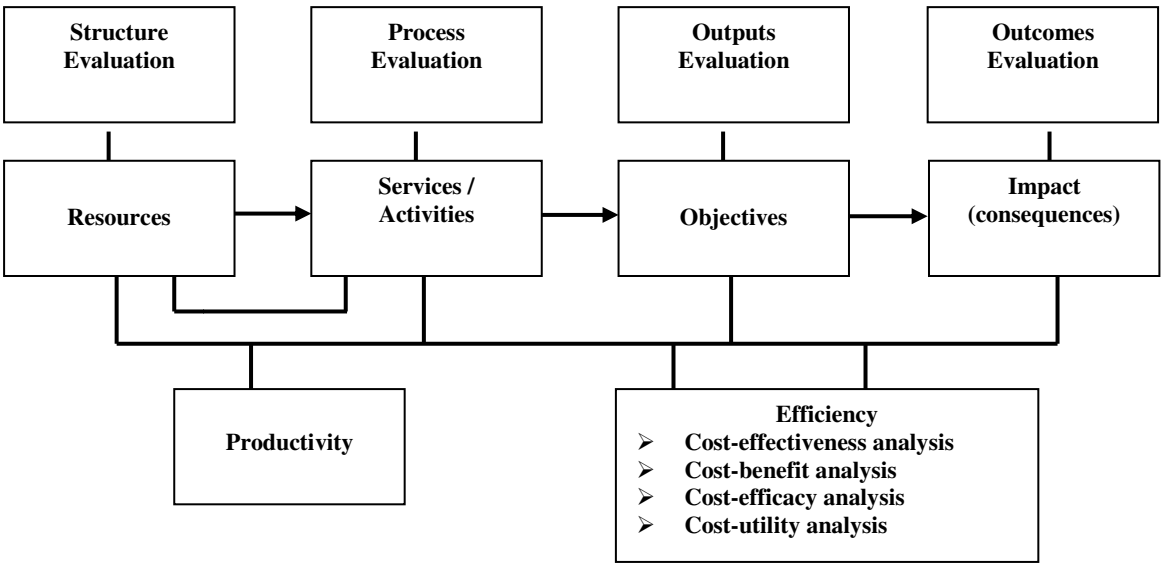
A more comprehensive evaluation (evaluative research) can also be done during this phase. It takes into account the relationships between the three components: structure, process, outputs/outcomes. For instance, a relation between different types/quantity of used resources can be estimated according to process or effects (outputs/outcomes).

Economic evaluation is the most appropriate tool for this purpose. There are two types of economic evaluation:

- Productivity analysis – establish a relation between the process (provided services/activities) and the resources used by the project (expressed as number of services per invested monetary unit, number of services per health professional etc.)
- Efficiency analysis – establish a relation between effects (output/outcome) and the resources used or provided services (both expressed in a monetary value)

A general framework of economic analysis was presented by R.Pineault (4) and is presented in figure 8.

Fig. 8 Economic analysis framework



Source: adapted after R. Pineault

The most important document of the evaluation is included in the Final Report. This document usually describes the successes and failures of the project. The content depends on the project nature. The content will generally focus on expected results versus achieved results, as well as on the short-term and long-term impact on the target population.

The achieved results can be grouped as follows:

- Physical results (degree of needs attainment reported to a reference status, the achieved level of indicators as a consequence of project implementation);
- Socio-cultural results (related to the improvement of quality of life, of the general health status, etc);
- Financial and economic results (reduction of sickness rate for the active population, etc);
- Non-measurable results (organizational change, capacity building, behaviour change etc).

The final report will also describe the degree of goal/objectives achievement, the quality of norms and standards used by the project, procedures and criteria requested by the financing agencies, the quality of collected information. The conclusions will outline the encountered difficulties and, if possible, their generating causes, and will make recommendations on results dissemination. This document represents a valid basis for policy-making.

Exercises

Project Development

Students will be provided with the following model for a project proposal. After each presentation during the lectures, the students will have to prepare every chapter of the project proposal according to the model. At the end of the module, each student will present his/her draft of project proposal.

Model for project proposal

1. **Project name**
2. **Executive summary**

(Brief statement of the problem, short description of the solution, funding requirements, brief description of the organization and its expertise).

3. **Background**

(Describe the context in which the project is developed; its relationship with other projects.)

4. **Project justification**

(Brief description of the problem that requires the project. Facts and statistics will be presented in annexes. Show how the project would contribute to the problem solving or reduction and what would be the consequences in case that the project will not be done.)

5. **Geographical coverage and target population**
6. **Project description**

- Goal
- Objectives
- Action plan
- Detailed schedule (use a Gantt chart)
- Detailed budget

7. Expected results

8. Monitoring, evaluation and reporting

(Use indices as much as possible.)

9. Arguments to the success and possible riscs

(feasibility, sustainability, etc)

10. Supporting materials

(ANNEXES: full description of the organization, CVs for the team members, recommendation letters, articles, statistics, documents that could support project utility, feasibility and sustainability).

References

1. Burton C, Michael N. A Practical Guide for Project Management. London (UK): Kogan Page Limited; 1992. p. 4-6.
2. Lock D. Project Management. 6th edition, Hampshire (England): Gower Publishing Limited; 1996.
3. Chalkley P. Project Management - Avoiding the pitfalls!, handouts from a training workshop on Project Management provided by Crown Agents - Management Training Centre, Bucharest, 1996.
4. Pineault R, Daveluy C. Health Planning - concepts, methods, strategies. Revised Edition. Ottawa (Canada): Editions Nouvelles; 1995.
5. Marcu A, Marcu Gr. Guidelines for the Health Programmes Management. Institute of Public Health Bucharest Publications; 2000.
6. Champagne F, Contandriopoulos AP, Pineault R. A conceptual Framework for health programmes evaluation. Montreal (Canada): Rev.Epidemiologie et Sante 1985;33:173-181.
7. Willoughby N. Incredibly Easy Project Management. Victoria (Canada): NWM Ltd 1994.
8. Wan T, Ozcan Y. Introduction to Grant Writing. Handouts from a workshop held at AIHA Third Partnership Conference for Central and Eastern Europe, Bucharest (Romania); 1998.