COMPARISON OF TEN BIOCHEMICAL LABORATORY TESTS BEFORE AND AFTER TREATMENT BY STATINS

TATIANA ĎURČEKOVÁ[1], JÁN MOCÁK[1*], JÁN BALLA[2], GABRIELA GROMANOVÁ[2], KATARÍNA BORONOVÁ[1]

[1]Department of Chemistry, University of SS. Cyril and Methodius, J. Herdu 2, Trnava, SK-917 01, Slovak Republic (tatiana.durcekova@ucm.sk)
[2]Analytical-diagnostic laboratory, Kováčska 15, Prešov, SK-080 01, Slovak Republic (jan.balla@mail.t-com.sk)

Abstract: Results of 10 biochemical tests of 172 patient data (among them 84 men data and 88 women data, resp.) before and after administration of statins were thoroughly studied. All monitored patients are characterized by disorders of lipoprotein metabolism or other kind of dislipidaemia. The calculations were performed using four chemometrical methods facilitating quantification and visualization of the statin effect upon most important biochemical parameters, mainly lipid markers, and allowing classification of the patient blood samples taking into account whether the patient has been or has not been medicated by a statin drug.

Key words: cardiovascular risk, lipid markers, statins, ROC curves, ANOVA, KNN classification.

1. Introduction

At the present state of medicine the patients with Coronary Heart Disease (CHD) may require intensive drug therapy usually realized by hypolipidemic drugs like statins. The number of individuals who are candidates for lipid-lowering therapy by statins has continued to increase.

Statins are chemically and pharmacologically diverse class of drugs that lower cholesterol levels in people with or at risk of cardiovascular disease by inhibiting the enzyme HMG-CoA reductase, which is the rate-limiting enzyme of the cholesterol synthesis (WIERBZBICKI *et al.*, 2003; LAWS *et al.*, 2004; KINLAY *et al.*, 2005; DEAMBROSIS *et al.*, 2009). Statins, the most widely prescribed medications in patients with hyperlipidaemia and coronary heart disease, have a number of pleiotropic actions beyond cholesterol lowering. They improve endothelial function, they have antioxidant and anti-inflammatory effects, they regulate neovascularization and have immunomodulatory activities (KWAK *et al.*, 2003; TOUSOULIS *et al.*, 2007; BLUM *et al.*, 2009).

The first part of the results dealing with most important and frequent biochemical tests of 172 patients (among them 84 men and 88 women, respectively) before and after administration of statins and processed by basic chemometrical procedures were given in our previous paper (DURCEKOVA *et al.*, 2009). In this work, further techniques of laboratory data processing were applied to inspect a diverse influence of the statin administration upon blood serum levels of the selected biochemical parameters. Several statistical/chemometrical tools have been here newly applied: (a)

the box-plot representation of the statin drug effect, (b) the comparison of efficiency of the selected variables (biochemical tests plus patient's age) by means of receiver operating characteristic curves (ROC) analysis and Gini indices, (c) one-way analysis of variance (ANOVA) demonstrating whether the selected variable is significantly changed with the statin administration, (d) the K-th nearest neighbour (KNN) pattern recognition technique used for classification of the patient samples before and after statin administration. This, rather complex approach allows investigate the statin effects on the patient's status from a different point of view.

2. Material and methods

2.1 Description of the analyzed biochemical data

Serum level of the selected biochemical tests was analytically determined for 172 probands (individuals) who were characterized by failure of their lipoprotein metabolism or another kind of lipidaemia. Two data matrices were investigated containing the test results for: (1) 84 *men* samples, and (2) 88 *women* samples, which represent the rows of the investigated data matrix. Its columns contained 12 monitored variables, namely 5 lipid parameters - total cholesterol (tCHOL), high density lipoprotein cholesterol (HDLc), low density lipoprotein cholesterol (LDLc), triacylglycerols (TG), aterogenity index (AI) – obtained by the following calculation: (tCHOL – HDLc)/HDLc, 6 standard biochemical parameters - creatinine (CREA), aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), creatinine kinase (CK), and gamma-glutamyl transferase (GGT), and, finally, the age of the patient as the last variable. All mentioned variables were measured before statin administration and one year after it.

2.2 Multidimensional data analysis

Multidimensional data analysis was performed using three software packages: (1) Minitab 15 was utilized to perform box-plots, (2) SPSS 15.0 was used to perform ROC analysis and ANOVA and (3) KNN was carried out using software package SAS Enterprise Guide 3.0.

3. Results and discussion

3.1 Comparison of selected biochemical parameters using box-plots

Box-plots (or box-and-whisker plots) are useful for the comparison of different groups of one-dimensional data. Its construction allows a visual representation of the data (MASSART *et al.*, 1997). The box itself covers inner 50 % of all data values, starting from the lower quartile (25 % of the ordered data) and ending by the upper quartile (75 % of the ordered data); the whiskers, represented by the abcisses, cover the lowest and highest part of the data representing the given variable. The line across the box represents the median.

Based on known categorization of probands, divided into two categories - before and after administration of statins, a comparison of the blood serum levels of the

studied variable before and after administration of statins was visualized using box-and-whisker plots and demonstrated for all studied variables. In addition, some multicomponent variables, namely PC1 (the first principal component), DF1 (the first discriminate function) and logit (the dependent variable in logistic regression) are also shown and compared to individual biochemical tests. The mentioned multicomponent variables were calculated by linear combinations of all original variables by principal component analysis, discriminant analysis and logistic regression (DURCEKOVA *et al.*, 2009). Further details of these multivariate techniques can be found in KHATTREE *et al.*, 2000; JOLLIFFE *et al.*, 2002; KLEINBAUM *et al.,* 2005. It is demonstrated in Fig. 1 that the selected two categories of probands are relatively well separated in case of variables LDLc, tCHOL and AI for the men samples and LDLc and tCHOL for the women samples. Nevertheless, it is evident that the best separation of two categories is achieved when using multicomponent variables DF1 and logit for the men samples and DF1, logit and PC1 for the women samples. The effective utilization of the multicomponent variables in prediction and/or confirmation of clinical diagnosis was discovered in our previous works (BALLA *et al.*, 2004; MRAZOVA *et al.*, 2009). A minor difference of HDL cholesterol between treated and untreated patients were expected. The change in box-plots of six standard biochemical parameters with the statin treatment was not significant.

3.2 ROC analysis

The predictive value of an assay can be displayed in the form of a plot of *Sensitivity* against (1 - *Specificity*), where Sensitivity and Specificity are defined in the way common in clinical chemistry (MOCAK *et al.*, 2003). The area under the corresponding curve, *A*, is defined between 0.5 (the diagonal line on the ROC diagram represents an assay with a zero predictive value and corresponds to 50 % of the total area) and 1.0 (absolutely successful laboratory test corresponds to the total area) (MASSART *et al.*, 1997; BALLA *et al.*, 2004). In the ROC analysis, the original laboratory variables as well as the multicomponent variables PC1, DF1 and logit were used. The variables with the largest ROC curve area are shown in Fig. 2, which demonstrates that the areas belonging to logit and DF1 are clearly larger compared to the best original variables (tCHOL, LDLc, AI). Among the original variables, the most significant statin effect exhibit tCHOL, LDLc, AI, TG and HDLc for the men samples and tCHOL, LDLc, AI and TG for the women samples with the ROC area greater that 0.6. The observed independence of six standard biochemical parameters characterizing liver and/or renal function upon statin treatment was anticipated and confirmed. An alternative, new-fashioned comparison of the statin effect upon 11 investigated variables is surveyed in Table 1 containing Gini coefficients, *G*, defined by means of the ROC curve area, *A*, in the more convenient interval (0, 1):

$$G = 2\,A - 1 \tag{1}$$

It is worth noting that negative values of Gini coefficients are insignificant since they were caused by calculation errors and denote in fact zero.
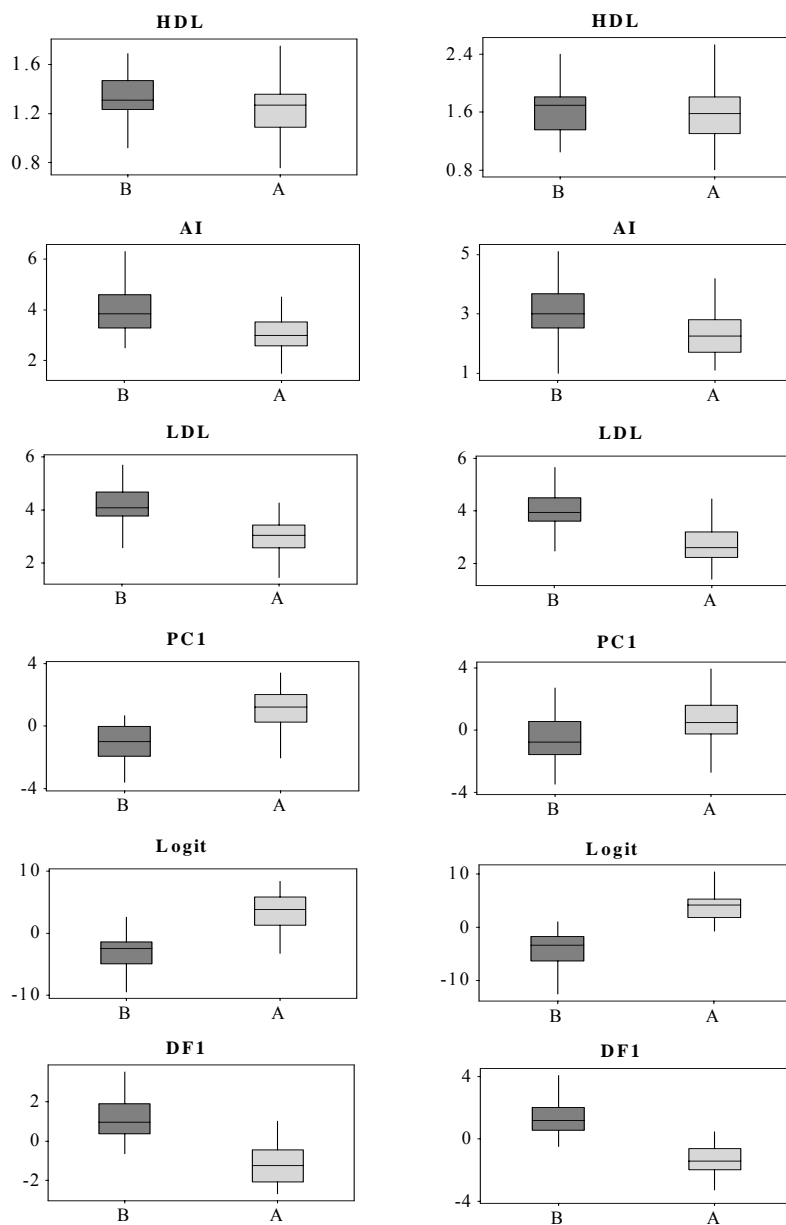
Fig. 1. Comparison of the box-and-whisker plots for the best laboratory tests and multicomponent variables DF1, PC1 and logit by box-and-whisker plots constructed using 10%, 25%, 50%, 75% and 90% percentiles of the ranked variable. B – blood serum samples before statin uptake, A – after the uptake. Men samples are located in the left frame, women samples are in the right frame. Software Minitab 15.1.0.0.
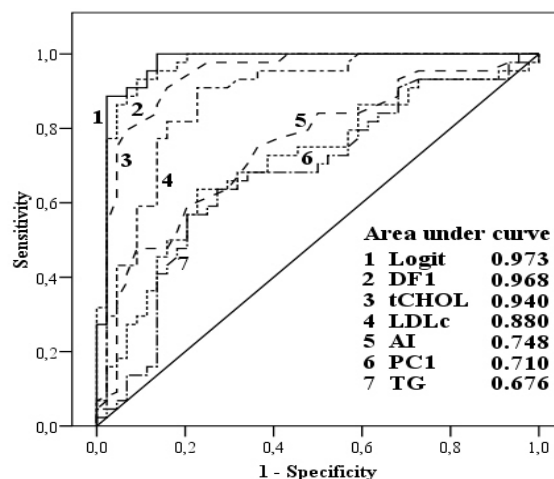
Fig. 2. ROC curves for the best biochemical parameters tCHOL, LDLc, AI and TG, and multicomponent variables DF1, Logit and PC1 selected by the largest area under the curve. 88 women samples. Software SPSS 15.0.

Table 1. Effect of statin administration on all investigated variables (laboratory as well as multicomponent ones) expressed by Gini coefficients.

Men			Women		
Variable	**A**	**G**	**Variable**	**A**	**G**
DF1	0.949	0.898	Logit	0.973	0.946
Logit	0.935	0.870	DF1	0.968	0.936
tCHOL	0.931	0.862	tCHOL	0.940	0.880
LDLc	0.921	0.842	LDLc	0.880	0.760
PC1	0.892	0.784	AI	0.748	0.496
AI	0.798	0.596	PC1	0.710	0.420
TG	0.713	0.426	TG	0.676	0.352
HDLc	0.629	0.258	HDLc	0.562	0.124
CREA	0.529	0.058	ALT	0.598	0.196
ALT	0.512	0.024	ALP	0.533	0.066
AST	0.509	0.018	GMT	0.517	0.034
ALP	0.501	0.002	AST	0.512	0.024
GMT	0.490	-0.020	CREA	0.509	0.018
CK	0.446	-0.108	CK	0.443	-0.114

3.3 Analysis of variance

The goal of analysis of variance, ANOVA, is to divide the total variance of the given variable into their appropriate components in order to discover whether the variable is significantly affected by the selected factor, which represents the outcome of the investigated problem (MASSART *et al.*, 1997). This factor is represented by a categorical variable, named here as *StatMW* and indicating whether the treatment by statins was carried out (after drug administration) or not (before) and if the patient is a

man or a woman: 1 – men before drug administration, 2 – men after the drug treatment, 3 – women before the treatment, 4 – women after the treatment. Two post-hoc ANOVA tests were applied: the least significant difference (LSD) test found out the variables tCHOL, LDLc, TG and AI capable to separate the categories 1 from 2 and 3 from 4 (which is the main goal), eventually also 1 from 4 and 3 from 2; Bonferroni test provided the same results. The selection of the most important ANOVA outputs is summarized in Table 2 where only the significant categories combinations are included.

Table 2. ANOVA outputs showing only those variables and categories I and J, which are significantly influenced by the statin treatment (with the significance *p*-value less than 0.05).

Multiple comparison							
Dependent variables	(I)	(J)	Mean Difference (I-J)	Stand. error	Sig.	95% confidence interval	
						Lower Bound	Upper Bound
tCHOL	1	2	1.567	0.172	0.000	1.228	1.905
	1	4	1.746	0.170	0.000	1.411	2.081
	3	2	1.489	0.170	0.000	1.154	1.824
	3	4	1.668	0.168	0.000	1.337	1.999
LDLc	1	2	1.243	0.156	0.000	0.936	1.550
	1	4	1.431	0.154	0.000	1.128	1.735
	3	2	1.061	0.154	0.000	0.758	1.365
	3	4	1.250	0.152	0.000	0.950	1.550
TG	1	2	0.555	0.169	0.001	0.226	0.888
	1	3	0.478	0.167	0.005	0.150	0.807
	1	4	0.879	0.167	0.000	0.551	1.208
	3	1	-0.478	0.167	0.005	-0.817	-0.150
	3	4	0.401	0.165	0.016	0.076	0.726
AI	1	2	0.908	0.179	0.000	0.555	1.260
	1	3	0.916	0.177	0.000	0.567	1.265
	1	4	1.657	0.177	0.000	1.308	2.006
	3	1	-0.916	0.177	0.000	-1.265	-0.567
	3	4	0.741	0.175	0.000	0.396	1.086

Significance level is expressed by *p*-values rounded to 3 decimal figures; the mean difference is considered significant when $p < 0.050$. Categories: 1 – men before drug administration, 2 – men after drug administration, 3 – women before drug administration, 4 – women after drug administration. Software SPSS 15.0.

3.4 KNN classification

K-th nearest neighbour, KNN, is a non parametric discrimination technique (KHATTREE and NAIK, 2000) very useful for classification purposes, which does not need any assumption on the distribution of errors. The main variant of this technique is based on the majority vote rule, which means that *K* neighbour objects, nearest to the classified object, are searched and then the classification of the given object is made according to which class the neighbour objects are predominantly classified.

KNN results were evaluated by the successful classification in % for (a) the training set, from which the classification model is calculated, and for (b) the samples leaved out from the training set by leave-one-out method, which is most important for prediction purposes. It is shown in Table 3 that the best classification performance was achieved when using seven nearest neighbours ($K = 7$) for the men samples and eleven for the woman samples.

Table 3. Classification results by KNN method; $K = 5 – 11$.

	K	**Results**	**Model**	**Leave-1-out**
M	5	true/all	69/84	64/84
		% true	**82.1**	**76.2**
	7	true/all	69/84	66/84
		% true	**82.1**	**78.6**
	9	true/all	70/84	64/84
		% true	**83.3**	**76.2**
	11	true/all	69/84	63/84
		% true	**82.1**	**75**
W	5	true/all	76/88	70/88
		% true	**86.4**	**79.5**
	7	true/all	75/88	73/88
		% true	**85.2**	**82.9**
	9	true/all	76/88	73/88
		% true	**86.4**	**82.9**
	11	true/all	76/88	75/88
		% true	**86.4**	**85.2**

M – men samples, W – women samples

4. Conclusions

Positive changes in lipid metabolism after statin treatment of the patients with cardiovascular risk can be unambiguously discovered and monitored by means of statistical and chemometrical techniques, which provide qualitative as well as quantitative judgment, which laboratory parameters are mostly affected by the administration of statin drugs. In this work, box-and-whisker plots and ROC curves allowed visualization of the statin effects. KNN classification method allowed a clear discrimination of the patients' samples into two categories – before and after statin treatment. Analysis of variance revealed that four variables are capable to differentiate the statin treatment with regard to the patient gender: aterogenity index, low density lipoprotein cholesterol, total cholesterol and triacylglycerols. High density lipoprotein cholesterol as well as all further investigated biochemical parameters were not efficient in differentiating neither men nor women groups before and after statin treatment.

A special feature of the achieved results is a very high diagnostic effectiveness of the multicomponent variables composed by linear combination of individual laboratory assays, which predestinates their further utilization in cardiovascular risk confirmation and prediction.

There may be considered two potentially serious side effects of statins of which patients need to be aware. Occasionally, statin use cause an increase in liver enzymes. If the increase is severe, the patient may need to stop taking the drug, which usually

reverses the problem. If there is no increase or it is only mild, one can continue to take the drug. In general, our study has not proved a significant change in the level of liver enzymes with the statin uptake. Such an effect should be individually assessed and perhaps further investigation should be made aimed to evaluating the mentioned side effects.

References

BALLA B., MOCAK J., PIVOVARNIKOVA H., BALLA J.: Comparative study of cardiovascular markers data by various techniques of multivariate analysis. Chemom. Intell. Lab. Sys., 72, 2004, 259-267.

BLUM A., SHAMBUREK R.: The pleiotropic effect of statins on endothelial function, vascular inflammation, immunomodulation and thrombogenesis. Atherosclerosis, 203, 2009, 325-330.

DEAMBROSIS P., TERRAZZANI G., WALLEY T., BADER G., GIUSTI P., DEBETTO P., CHINELLATO A.: Benefit of statins in daily practise? A six-year retrospective. Pharmacol. Res., 2009 (in press).

DURCEKOVA T., MOCAK J., BALLA J., GROMANOVA G., BORONOVA K.: Relationship between administration of statins and blood serum levels of selected biochemical parameters. Nova Biotechnol., 9, 2009, 83-90.

LAWS P.E., SPARK J.I., COWLED P.A., FITRIDGE R.A.: The role of statins in vascular disease. Eur. J. Vasc. Endovasc. Surg., 27, 2004, 6-16.

JOLLIFFE I.T.: Principal component analysis, Springer-Verlag, New York, 2002, 487 pp.

KINLAY S.: Potential vascular benefits of statins. Am. J. Med., 118, 2005, 625-675.

KLEINBAUM D.G., KLEIN M., PRYOR E.R.: Logistic regression, Springer, Heidelberg, 2005, 282 pp.

KHATTREE R., NAIK D.N.: Multivariate Data Reduction and Discrimination. SAS Institute, Cary, North Carolina, 2000, 558 pp.

KWAK B., MULHAUPT F., MACH F.: Atherosclerosis: anti-inflammatory and immunomodulatory activities of statins. Autoimmun. Rev., 2, 2003, 332-338.

MASSART, D.L., VANDEGINSTE, B., BUYDENS, L., De JONG, S., LEWI, P., SMEYERS-VERBEKE, J.: Handbook of Chemometrics and Qualimetrics, Part A, Elsevier, Amsterdam, 1997, 886 pp.

MOCAK J., BALLA B., BOBROWSKI A., BLAZICEK P.: Proper ways of comparison of two laboratory methods. Chem. Pap., 57, 2003, 143-146.

MRAZOVA V., MOCAK J., VARMUSOVA E., KAVKOVA D., BEDNAROVA A.: Use of multidimensional data analysis for prediction of lung malignity. J. Pharm. Biomed. Anal., 50, 2009, 210-215.

TOUSOULIS D., CHARAKIDA M., STEFANADI E., SIASOS G., LATSIOS G., STEFANADIS C.: Statins in heart failure. Beyond the lipid lowering effect. Int. J. Cardiol., 115, 2007, 144-150.

WIERZBICKI A.S., POSTON R., FERRO A.: The lipid and non-lipid effects of statins. Pharmacol. Therapeut., 99, 2003, 95-112.