Journal of Pharmacy

ORIGINAL ARTICLE

Open Access

Validity of claims database compared with the electronic medical record of private health clinics in Malaysia: A pilot study

Nor Hidayah Mohd Taufek[1], Norny Syafinaz Ab Rahman[1], Nor Ilyani Mohamed Nazar[1], Siti Halimah Bux[1], Che Suraya Zin[1*]

ABSTRACT

Introduction: The validity of health insurance claims data in private health sectors has been widely reported in many developed countries to monitor details of private healthcare utilisation. Little is known regarding the data validity of private health care services and insurance claims in Malaysia. This pilot study aimed to validate the claims data from a private health insurance database, using electronic medical records (EMR) at the private clinics as the gold standard.

Method: Patients' data were retrieved from the PMCare health insurance database from 2016-2019 recorded for International Islamic University Malaysia employees. Patients were sampled from the PMCare database and manually compared with data from EMR of selected private panel clinics. Data were analysed for descriptive statistics using Microsoft Excel 2013.

Results: A total of four panel clinics consented to the study, and data were available for 2016, 2017 and 2019. The number of observations obtained from 118 patients (male = 63, female = 55) was 386, with the most common diagnosis reported in the PMCare database was acute upper respiratory tract infection (63.6%). Total accuracy between PMCare and EMR data was 91.5%, with an 8.5% difference or inaccuracy. Percentage accuracy was varied between different clinics (A= 92.6%, B=84.7%, C=98.6%, D=82.6%).

Conclusion: Data submitted to PMCare claims by private health clinics had high accuracy (>90%) and is acceptable for research and other applications. Future studies should investigate the differences in clinic-based practice for documenting the identified types of discrepancies to improve the accuracy of private health insurance databases.

ARTICLE HISTORY:

Received: 20 August 2021
Accepted: 31 January 2022
Published: 31 January 2022

KEYWORDS:

Validity, claim database, electronic medical record, private health clinics, accuracy.

HOW TO CITE THIS ARTICLE:

Mohd Taufek, N. H., Ab Rahman, N. S., Mohamed Nazar, N. I. Bux, S. H. & Zin C. S. (2022). Validity of claims database compared with the electronic medical record of private health clinics in Malaysia: A pilot study. *Journal of Pharmacy*, 2(1), 31-35.

doi: 10.31436/jop.v2i1.107

***Corresponding author:**

Email address: chesuraya@iium.edu.my

Authors' Affiliation:

[1] *Big Data Research Group, Department of Pharmacy Practice, Kulliyyah of Pharmacy, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia*

Introduction

The healthcare system in Malaysia has been primarily divided into private and public sectors. High concentrations of private practices utilised by the community are covered by private health insurance companies (Thomas et al., 2011). Many private practices have also been affiliated as panel health clinics to government and private organisations which subsidise certain medical insurance coverage as health benefits for their employees with variable premiums charges (Chua &Cheah, 2012). Private health insurance companies record the data on patients' claims such as diagnosis, prescribed medications and cost submitted by the panel health clinics from their medical records. The format for the claims database could be different between companies, but it generally collects information on prescriptions, procedures and administrative costs of inpatients and outpatients.

Claims data are useful to monitor details regarding healthcare utilisation, including disease and prescription patterns. However, it requires high accuracy with good validity and reliability components to be tested so that the characteristics and limitations of the dataset are well understood (Iwamoto et al., 2015; Du et al., 2006). Validity studies are crucial owing to increased reliance on health insurance claims, and thus the credibility of such data should be confirmed (Koram et al., 2019). There have been limited validation studies reported in the Asia Pacific region, including Australia, South Korea, New Zealand, Thailand, Singapore and Japan (Koram et al., 2019). To our knowledge, there have been no studies available reporting the validity of the claims database in Malaysia. Little is known whether prescription patterns, medication utilisation, and cost can be accurately obtained from private insurance databases. It is important to test the accuracy of insurance database documentation to ensure the credibility of research utilising databases. Validity of database of private insurance companies will facilitate research towards sustainable reimbursement policy, optimal utilisation of medications and patient outcomes.

A review has highlighted that the gold standard for database validity research was obtained from medical records, registry data, self-reported questionnaires and other data sources (Koram et al., 2019). Validated outcomes included variables such as medical conditions (Takeda et al., 2016), risk monitoring (Tomlin, Reith & Woods, 2017) and disease-specific comorbidities (Yamana et al., 2017). The present study aimed to investigate the accuracy of the claims data of a private health insurance company, PMCare, compared with electronic medical records (EMR) of private panel clinics that are covered by PMCare.

Methodology

This study was approved by the International Islamic University Malaysia Ethical Committee (IREC-2019-212). It was conducted from February to July 2020 at panel private health clinics in two states of Pahang and Selangor, Malaysia. The results were reported in an aggregated manner using de-identified data. Informed consent was not required as this study did not involve direct patient interaction.

This retrospective pilot study collected data for 2016-2019 provided by the private health insurance company PMCare for outpatient settings. The PMCare insurance was subsidised by International Islamic University Malaysia (IIUM) for its employees. Based on the policy, the claim limit was set at Ringgit Malaysia (RM) 45 per visit per patient. The panel private health clinics (PHC) were identified to collect the original data from EMR. The inclusion criteria were clinics with a high frequency of attendance by IIUM employees, used EMR and consented to the study. The exclusion criteria were those with low frequency of patients, used manual medical records and did not consent to the study. A total of four panel clinics were included, and data were collected from EMR.

Patient selection

The selection of patients was conducted using systematic sampling. Three patients per month were selected at the beginning, middle and end of each month from the PMCare database. The inclusion criteria were patients who had three or more medications. The search was conducted using a unique claim code that extracted the correct data regardless of the date recorded for database claim, patient's visit or delayed submission of claims to PMCare. Data collection included information on claim code, date of prescription, name of the patient, patient identification, diagnosis, name of medications, number of medications per prescription, the number of drugs supplied, consultation fee, the total cost of medications and amount paid by the patient. Validated outcomes included diagnosis, medication name, number of medications per prescription, the quantity of medication, and the amount claimed to the health insurance. Majority of previous studies (83%) have reported similar data collection methods where cases were most frequently validated against diagnoses in medical records (Widdifield et al., 2013).

Accuracy measurement

The accuracy was obtained by comparing the PMCare dataset with the original data from EMR by dividing the number of observations found to be the same by the total number of data elements of the PMCare database then multiplying that total by 100. The agreement exceeding 90% was considered as a good benchmark when data elements from the two sources were found not to be the same, noting the reasons (for example, unclear data definitions) as reported in previous studies (Iwamoto et al., 2015; Miller et al., 2009). The discrepancy was determined when the inaccuracies of observations recorded in the EMR and PMCare database were identified. The discrepancies were calculated by dividing the number of observations found to be different with the total observations from selected PMCare data then multiplying by 100.

Results

A total of 118 patients (age: mean ±SD, 33.2 ± 12.7) were selected based on the availability of data from EMR (Table 1). Although there were three patients selected per month from 2016-to 2019, many patients had to be excluded due to changes in the recording system used at the clinics, crash of the software and incomplete record of EMR. We found that several observations from EMR recorded extra medications (415 observations) compared to the PMCare database, which was 386 observations. Electronic medical records contained original data entered from patients' prescriptions. The number of observations recorded in the EMR was higher than the PMCare database because the EMR captured the total number of medications prescribed to each patient.

In contrast, only a few medications were recorded for insurance claims to PMCare for the coverage limit of RM45. Therefore, after removing extra observations from EMR for accuracy measurement, the total number of observations for both EMR and PMcare database was 386. As presented in Table 2, the number of patients included in the present pilot study was sufficient as reported by a similar study that measured accuracy and validity of a hospital database (Cook et al., 2002).

The advantage of using EMR was that additional information was available and comprehensively recorded for each visit. We found that consultation fees were varied between clinics and between clients depending on the number of medications prescribed and adjustment to the insurance eligibility. Data from EMR provided the amount of out-of-pocket money paid by the patients if the total charges for consultation and medications exceeded the insurance coverage of RM 45, including the detailed price for each medication with a specific quantity. On the contrary, the PMCare database did not have this information. Discrepancies were found due to differences in recording the name and quantity of medications, diagnosis, the number of medications and the missing information in either EMR or PMCare database for the variables investigated. The most common diagnosis reported in the PMCare database was acute upper respiratory tract infection (URTI) (63.6%), followed by acute gastroenteritis and colitis (13.5%).

Table 1: Demographics of patients from four private medical clinics.

Characteristics	N = 118	Frequency (%)
Gender:		
a. Male	63	53.4
b. Female	55	46.6
Age (years):		
Mean (SD)	33.2 (12.7)	
Median (range)	35 (4 – 61)	
Diagnosis:		
a. Acute Upper Respiratory Infection, unspecified	75	63.6
b. Infectious Gastroenteritis and Colitis, Unspecified	16	13.5
c. Unspecified Abdominal Pain	3	2.54
d. Anemia, Unspecified		
e. Others (Urticaria, Low back pain, Acute Tonsillitis, Viral Infection Unspecified, Dermatitis and Eczema, Conjunctivitis, Gout, Hemorrhoids, Hordeolum (Externum) (Internum) of Eyelid, Gastritis, Myalgia, Mycosis)	3 21	2.54 17.8

Discussion

The accuracy of the PMCare database has never been validated to determine its credibility to be used for database research in Malaysia. The present study reported that the PMCare claims database had overall 91.5% accuracy and 8.5% discrepancies compared with EMR data as a gold standard for four different panel clinics for the segregated patient samples collected in 2016, 2017 and 2019. This outcome indicated an ideal accuracy value that was set as above 90% to be considered valid (Miller et al., 2009). Our finding was lower than that reported by a study in Japan, that demonstrated 99% accuracy in estimating opioid consumption among cancer patients (Iwamoto et al., 2015).

Another study validated claims algorithm and clinical operative reports with 97% accuracy (Miller, Saigal & Warren, 2009). However, a different study has indicated a lower range of good accuracy with 80% and above for a pilot study reporting database validity (Cook et al., 2002). The analysis for an individual clinic showed that accuracy could achieve above 98% with a systematic and appropriate recording of claims data according to EMR (Table 2).

Table 2: Discrepancies of data from EMR of respective clinics in comparison to PMCare Database

Variables	Clinic A	Clinic B	Clinic C	Clinic D	Total
No. of patients	30	15	50	23	118
Year of observations	2016	2016	2016-2017	2019	
No. of Observations from EMR	94	59	147	86	386
No. of Observations from PMCare claims database	94	59	147	86	386
No. of observations with discrepancies	7	9	2	15	33
Types of discrepancies	Name and quantity of medication.	Name and quantity of medication, diagnosis.	Name of medication.	Name and quantity of medication, diagnosis.	
Percentage of discrepancies (%)	7.4	15.3	1.4	17.4	8.5
Percentage of accuracy (%)	92.6	84.7	98.6	82.6	91.5

Investigation of data discrepancy showed that types of discrepancy tend to vary between clinics. This could be due to the nature of PMCare claims data, where data were created and submitted later from the date of the patient's visit, whereas the EMR data showed the actual medications dispensed daily. Although the differences in characteristics exist, both data were prone to human error in entering the details into the online system, which resulted in discrepancies (Iwamoto et al., 2015). It is also important to note that claims databases could only cater for specific data covered by the insurance company for certain medications and diseases. A review of previous studies has also highlighted the critical measures of diagnostic accuracy (sensitivity, specificity, and positive and negative predictive values) and disease prevalence in reporting validity of administrative database algorithms (Widdifield et al., 2013).

Our study identified the discrepancies by including the missing names of medication or different medications recorded when compared between EMR and PMCare database. The diagnosis error was included when the diagnosis submitted to the PMCare database was different or unspecified compared to EMR. The limitation of this study was that we did not have specific details regarding the options given for each data category in the claim database system, such as general option (e.g. unspecified abdominal pain) or specific option for accurate diagnosis (e.g. dyspepsia, gastritis). We did not consider that specific, or general diagnosis of similar nature or location was different unless the difference was obvious such as abdominal pain (EMR) versus URTI (PMCare). From the perspective of the claims database and EMR, we propose that details listed in the claims database should be the same as those recorded in EMR. It was expected that the number of observations in EMR would be greater than those in the claim database since the insurance company covers only certain medications. The consistency in recording the data will improve the accuracy and validity of the data from both sources.

The major limitation of our pilot study was that the sample size was relatively small compared to previous studies (Iwamoto et al., 2015; Takeda et al., 2016) but was sufficient to report validity for a pilot study (Cook et al., 2002). Our study was multi-centred, with a standardised data format from the PMCare database, which allowed the comparison of varying practices between settings. Future studies may identify interventions for corrective actions to ensure high accuracy of the claims database. The results suggest that PMCare database is valid and accurate as a valuable resource to be used for research purposes but with careful analysis and interpretation considering the data characteristics. The identified discrepancies were common for certain variables in database validity studies as previously reported in other studies (Cook et al., 2002; Gabriel et al., 2005). However, improved documentation practice is required at the specific settings to improve the accuracy of the claims database.

Conclusion

Data from the health insurance claims database was valid and applicable for research purposes with careful analysis and interpretation. The overall high accuracy of the health insurance claims database in this study indicated its validity. However, variation in accuracy level was found between different clinics attributable to varying practices in recording the data. Future studies are required to identify effective interventions to improve the accuracy of claims databases with gold standards.

Acknowledgement

The authors would like to thank the PMCare database health insurance company and private panel health clinics of the International Islamic University Malaysia for their assistance in completing this research. CSZ was supported by a research grant from IIUM Health Wellness Centre.

Data Availability Statement

Request to access the datasets should be directed to International Islamic University Malaysia Research Ethical Committee. The de-identified data could be shared with interested researchers after obtaining approval from the above ethical committee. The restriction on public data deposition is due to the privacy and confidentiality of patients' health data.

Conflict of Interest

The authors have no conflict of interest to declare. The funder had no role in study design, data collection and analysis, decision to publish or prepare the manuscript.

References

Chua, H.T., & Cheah J.C. H. (2012). Financing universal coverage in Malaysia: a case study. BMC Public Health, 12(Suppl 1), S7. https://doi.org/10.1186/1471-2458-12-S1-S7

Cook, S.F., Visscher, W.A., Hobbs, C.L., & Williams R.L. (2002). The Project IMPACT Clinical Implementation Committee Project IMPACT: Results from a Pilot Validity Study of a New Observational Database, Critical Care Medicine, 30 (12), 2765-2770. doi: 10.1097/00003246-200212000-00024

Du, X.L., Key, C.R., Dickie, L., Darling, R., Geraci, J.M., & Zhang, D. (2006). External validation of Medicare claims for breast cancer chemotherapy compared with medical chart reviews. Med Care., 44(2), 124–131. doi: 10.1097/01.mlr.0000196978.34283.a6

Gabriel, S.E., Amadio, P., & Ilstrup, D. (2005). The Feasibility and Validity Studies Comparing Orthopedists and Non-Orthopedists caring for musculoskeletal Injuries: Results of a Pilot study. Health Sciences Research, 10(3), 163-168. https://doi.org/10.1002/art.1790100303

Iwamoto, M., Higashi, T., Miura, H., Kawaguchi, T., Tanaka, S., Yamashita, I., Yoshimoto, T., Yoshida S., & Matoba, M. (2015). Accuracy of using diagnosis procedure combination administrative claims data for estimating the amount of opioid consumption among cancer patients in Japan. Japanese Journal of Clinical Oncology, 45(11), 1036–1041. https://doi.org/10.1093/jjco/hyv130

Koram, N., Delgado, M., Stark, J.H., Setoguchi, S., & de Luise, C. (2019). Validation studies of claims data in the Asia-Pacific region: A comprehensive review. Pharmacoepidemiology and Drug Safety, 28(2), 156–170. https://doi.org/10.1002/pds.4616

Miller, D.C., Saigal, C.S., Warren, J.L., Leventhal, M., Deapen, D., Banerjee, M., Lai, J., Hanley, J., & Litwin, M.S. (2009). External validation of a claims-based algorithm for classifying kidney-cancer surgeries. BMC Health Services Research, 9 (92). https://doi.org/10.1186/1472-6963-9-92

Thomas, S., Beh, L., & Nordin, R.B. (2011). Health care delivery in Malaysia: changes, challenges and champions. Journal of Public Health in Africa, 2(2), e23. https://doi.org/10.4081/jphia.2011.e23

Tomlin, A.M., Reith, D.M., Woods, D.J., Lloyd, H.S., Smith, A., Fountain, J.S. & Tilyard, M.W. (2017). Pharmacoepidemiology database system for monitoring risk due to the use of medicines by New Zealand primary care patients. Drug Safety, 40(12), 1259-1277. https://doi.org/10.1007/s40264-017-0579-1

Takeda, T., Mihara N., Murata T., Shimai Y., Okada K., Manabe, S., & Matsumura, Y. (2016). Estimating the ratio of patients with a certain disease between hospitals for the allocation of patients to clinical trials using health insurance claims data in Japan. Studies in Health Technology and Informatics, 228, 537-541. Doi: 10.3233/978-1-61499-678-1-537

Widdifield, J., Labrecque, J., Lix, L., Paterson, J.M., Bernatsky, S., Tu, K., Ivers, N., Bombardier, C. (2013). Systematic review and critical appraisal of validation studies to identify rheumatic diseases in health administrative databases. Arthritis Care & Research, 65(9), 1490-1503. https://doi.org/10.1002/acr.21993

Yamana, H., Moriwaki, M., Horiguchi, H., Kodan, M., Fushimi, K., & Yasunaga, H. (2017). Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. Journal of Epidemiology, 10, 476–482. https://doi.org/10.1016/j.je.2016.09.009