Yonsei Medical Journal
Vol. 41, No. 5, pp. 570-576, 2000

The Accuracy of Myocardial Infarction Diagnosis in Medical Insurance Claims

So Yeon Ryu[1], Jong Ku Park[2], Il Suh[3], Sun Ha Jee[4], Jong Park[1], Chun-Bae Kim[2], and Ki Soon Kim[1]
For the Korean Research Group for Cardiovascular Disease Prevention and Control

Abstract

We attempted to assess the accuracy of the International Classification of Diseases (ICD) codes for myocardial infarction (MI) in medical insurance claims, and to investigate the reasons for any inaccuracy. This study was designed as a preliminary study to establish a surveillance system for cardiovascular diseases in Korea. A sample of 258 male patients who were diagnosed with MI from 1993 to 1997 was selected from the Korea Medical Insurance Corporation cohort (KMIC cohort: 183,461 people). The registered medical record administrators were trained in the survey technique, and gathered data by investigating the medical records of the study subjects from March 1999 to May 1999. The definition of MI for this study included symptoms pursuant to the diagnostic criteria of chest pain, electrocardiogram (ECG) findings, cardiac enzyme and results of coronary angiography or nuclear scan. We asked the record administrators for the reasons of incorrectness for cases where the final diagnosis was 'not MI'. The accuracy rate of the ICD codes for MI in medical insurance claims was 76.0% (196 cases) of the study sample, and 3.9% (ten cases) of the medical records were not available due to hospital closures, non-computerization or missing information. Nineteen cases (7.4%) were classified as insufficient due to insufficient records of chest pain, ECG findings, or cardiac enzymes. The major reason of inaccuracy in the disease code for MI in medical insurance claims was 'to meet the review criteria of medical insurance benefits (45.5%)'. The department responsible for the inaccuracy was the department of inspection for medical insurance benefit of the hospitals.

Key Words: Surveillance system, medical insurance claims, accuracy rate, myocardial infarction

INTRODUCTION

Since 1970, the occurrence of diseases in the Republic of Korea has shifted from acute infectious disease to chronic degenerative disease. Morbidity and mortality from cardiovascular diseases (CVD) have rapidly increased. According to the statistical analysis for the cause of death in Koreans in 1997, CVD was 23.4% of all registered death.[1] Regarding each kind of CVD, the mortality rate caused by a cerebrovascular accident (CVA) was the highest, and the rate by ischemic heart disease (IHD), which has risen five or six times in the past ten years, is expected to increase considerably and to be large in scale among the aged population.[2-4]

There is a decreasing mortality rate of CVD in developed countries, which seems to be the result of the investigation of risk factors through well-organized epidemiologic studies and the systematic administration of nationwide prevention and management projects on the investigated risk factors.[5,6] For example, there is a CVD surveillance system in the US that is currently being developed by the National Institute of Health (NIH) through a five-year project since 1998. Considering the deterioration in the quality of life, the enormous medical cost required for treatment and rehabilitation, and the death caused by CVD, it is necessary to provide countermeasures to decrease the death and the disease onset by immediately establishing a national surveillance system like those found in developed countries.[7]

Public health surveillance is an ongoing systematic

Received June 9, 2000
Accepted July 21, 2000
[1]Department of Preventive Medicine, College of Medicine, Chosun University, [2]Department of Preventive Medicine, Yonsei University Wonju College of Medicine, [3]Department of Preventive Medicine and Public Health, Yonsei University College of Medicine, [4]Graduate School of Health Science and Management, Yonsei University.
This study was supported by the research fund for health promotion of Ministry of Health and Welfare and Korea Research Foundation grants (1998-021-F00220) in 1998.
Address reprint requests to Dr. K. S. Kim, Department of Preventive Medicine, College of Medicine, Chosun University, Kwangju 501-759, Korea. Tel: 82-62-220-3674, Fax: 82-62-225-8293, E-mail: ksdkim@mail.chosun.ac.kr

collection, analysis, and interpretation of outcome-specific data for use in the planning, implementation, and evaluation of public health practice.[8] The most difficult and time-consuming part of disease surveillance is the data collection process. Certain general principles can be applied for collecting the least biased data. First, the informants should be motivated to supply the best information possible. Second, a simple survey and definite items are required. Third, the data should be collected on a timely basis. Last, a system should be set up to minimize the omission of reports about patients.

In the case of operating a surveillance system for CVD in Korea, the information related to the onset of the diseases can be available from hospital reports and the disease codes in medical insurance claims. To create accurate data, the system needs a connection between the hospital reports and the medical insurance data, but it doesn't seem to be as realistic as expected. Even though both data sources are virtually the same, because they are both based on the discharge diagnosis of the hospitals, there is a possibility of coding error related to the criteria for insurance benefits, misdiagnosis and duplicate admission.[9-14]

The purpose of this study was to assess the accuracy of the disease codes for myocardial infarction (MI) in medical insurance claims in order to learn how medical insurance claims can be used for acquiring data about MI. This study was designed as a preliminary study to establish a nationwide surveillance system for CVD.

MATERIALS AND METHODS

Study population

The Korea Medical Insurance Corporation (KMIC) provided health insurance to civil service workers, teachers, and their dependents in Korea. Of the entire Korean population (approximately 43 million in 1990), 4,603,361 (11%) were covered by KMIC, including 1,213,594 primary members and their 3,389,767 dependents. All insured workers were required to participate in biennial health screening examinations supported by KMIC. In 1990 and 1992, 95% and 94%, respectively, completed these biennial examinations.

The KMIC study cohort consisted of 115,600 men and 67,861 women, aged 35 to 59 years, who attended both the 1990 and 1992 health screening examinations. The male sample was chosen from a 25% systematic sample with a random start, drawn from insured members ordered by national identification number. The females included all insured members who met the same criteria in age and health screening as the males. Of the KMIC study cohort, 892 cases were hospitalized due to ischemic heart diseases (International Classification of Diseases, Ninth Revision [ICD-9][15] codes 410-414) from 1993 to 1997, so they were the population of this study (Table 1).

Selection of study subjects

The selection of study subjects from the population was based on the several criteria. First, the male patients with MI (ICD-9 code 410) were selected. Second, a random 70% of the hospitals with certified medical record administrators and all medical institutions without certified medical record administrators in Pusan and Seoul were selected. Third, if patients were diagnosed with the same disease by two or more hospitals or were admitted to the hospitals repeatedly within 28 days, they were classified as duplicate admission and only the medical record of the first admission was investigated.[16] In total, there

Table 1. Study Population According to the KMIC Cohort*

Unit : person

Year	1993	1994	1995	1996	1997	Total
Study population	108,802	103,776	100,018	97,266	94,562	
No. of inpatients	4,245	5,571	5,688	5,592	5,436	26,532
Ischemic heart disease patients	101	188	203	212	188	892

*Korea Medical Insurance Corporation Cohort: 115,600 people (male).

were 258 people in the study sample (Table 2).

Data collection

To execute this study smoothly, a research team from the Korean Research Group for Cardiovascular Disease Prevention and Control was organized with six medical colleges from all over the country. Our research was supported by the Korean Medical Record Association and the Ministry of Health and Welfare.

Table 2. Number of Study Hospitals (clinics) and the Study Sample

Medical record administrator	Registered		Unregistered	
Region	No. of hospitals	No. of MI	No. of hospitals/ clinics	No. of MI
Seoul	23	87	5	6
Pusan	9	32	3	3
Inchon · Kyonggi	8	17	–	–
Kangwon	4	13	–	–
Chungbuk	1	9	–	–
Taejeon · Chungnam	2	14	–	–
Chonbuk	3	14	–	–
Kwangju · Chonnam	3	22	–	–
Taegu · Kyungbuk	9	31	–	–
Wulsan · Kyungnam	2	10	–	–
Total	64	249	8	9

A survey sheet for MI was created to confirm the consistency between the disease codes in medical insurance claims and those in the discharge medical records for each subject. After reviewing the results of various MI-related international studies (e.g. ARIC (atherosclerosis risk in community) Study,[17] Cardiovascular Heart Study,[6] Minnesota study,[18] and the MONICA (multinational monitoring of trends and determinants in cardiovascular disease) study[19]), certain workshops and pilot studies were executed four times between February and March of 1999, and a survey sheet for the accuracy of MI diagnosis was finalized. The contents of the sheet were composed of age, date of onset, survival at discharge, diagnostic data (chest pain, level of cardiac enzymes and ECG findings), previous MI history, and clinical diagnosis code by the ICD.[15]

Prior to collecting the data, the medical record administrators at selected hospitals were educated about the importance of the survey and how to gather data with written survey protocol by investigators on six regional research teams. After educating the 69 medical record administrators working at 64 hospitals, the study subjects' medical records were reviewed and the survey was conducted between March 26 and May 21, 1999.

The survey sheets were transferred to regional research teams and the accuracy of disease diagnosis was assessed according to the diagnostic criteria for MI adopted for this study. The diagnostic criteria for MI (Table 3) followed the algorithm of the ARIC study[6,17] and the MONICA project.[18-20] According to

Table 3. Diagnostic Criteria of Myocardial Infarction

ECG pattern	Cardiac enzyme classification			
	Abnormal	Equivocal	Insufficient	Normal
Cardiac pain present (typical or atypical)				
Definite (diagnostic Q wave)	Definite	Definite	Definite	Definite
Definite (equivocal Q or ST−T abnormality)	Definite	Definite	Probable	No MI
Probable or ischemic	Definite	Probable	No MI	No MI
Uncodable or insufficient	Definite	No MI	No MI	No MI
Cardiac pain absent				
Definite (diagnostic Q wave)	Definite	Definite	Definite	Probable
Definite (equivocal Q or ST-T abnormality)	Definite	Probable	No MI	No MI
Probable or ischemic	Probable	No MI	No MI	No MI
Uncodable or insufficient	No MI	No MI	No MI	No MI

Source: ARIC cohort diagnostic criteria for hospitalized MI (1989).

the criteria, each MI was classified as "definite", "probable", "no MI" or "insufficient data". If the cases were diagnosed as MI by nuclear scanning or coronary angiography, they were classified as "definite".

Data analysis

To calculate the accuracy of the disease codes shown in the medical insurance claims, the diagnoses were judged as accurate when the cases were classified as "definite" or "probable" from the investigation of medical record by the diagnostic criteria. If the cases were "no MI", they were judged as inaccurate.

The accuracy could be calculated by two formulas according to the count of numerator. First, the accuracy among all the study subjects was calculated as a percentage by dividing the accurate MI cases (definite and probable) into total study subjects (Formula 1). Second, the accuracy among the investigated subjects was calculated as a percentage by dividing the accurate MI cases into the study subjects, excluding cases whose data were impossible to assess because of insufficient data, or were unavailable due to hospital closures, absence of medical records, etc. (Formula 2).

$$\frac{\text{No. of MI cases judged definite or probable}}{\text{No. of total study subjects}} \times 100 \cdots \text{Formula 1}$$

$$\frac{\text{No. of MI cases judged definite or probable}}{[\text{No. of total subjects - No. of cases whose data were impossible to assess due to insufficient data, or were unavailable due to hospital closures, absence of medical records, etc.}]} \times 100 \cdots \text{Formula 2}$$

Descriptive statistics (frequency, percent) were used to review the causes of inaccuracy of the diagnosis.

RESULTS

Accuracy of the myocardial infarction diagnosis

Of the total 258 subjects, the data for ten (3.9%) subjects were not available due to hospital closure, a non-computerized system, or missing medical records. Nineteen (7.4%) subjects were classified as insufficient cases due to insufficient records. One hundred ninety-six cases were found to be accurate diagnoses. The accuracy was 76.0% among all the study subjects by Formula 1, and it was 85.6% among the cases with available and assessable medical record data by Formula 2 (Table 4).

Reasons for inaccurate myocardial infarction diagnosis

There were 33 (12.8%) subjects whose cases were

Table 4. The Accuracy of Myocardial Infarction Codes in Medical Insurance Claims

Medical records: Availability	interpretation		Person (%)	(%)
Available	Possible	Accurate	196 (85.6)	(76.0)
		Inaccurate	33 (14.4)	(12.8)
		Subtotal	229 (100.0)	(88.8)
	Impossible*		19	(7.3)
Unavailable†			10	(3.9)
Total			258	(100.0)

* Due to insufficient data.
† Due to closure of medical facilities, absence of computer system, omission of medical records, etc.

Table 5. Reasons for the Inaccuracy for Myocardial Infarction Diagnosis

Main reasons	Case (%)
To meet review criteria of medical insurance benefits	15 (45.5)
Diagnostic error	2 (6.1)
Incorrect coding habits	1 (3.0)
Typing error	1 (3.0)
Unknown	14 (42.4)
Total	33 (100.0)

Table 6. Departments Making Inaccurate Myocardial Infarction Diagnoses

Main department	Case (%)
Dept. of inspection for medical insurance benefit	16 (48.5)
Medical part (doctors)	3 (9.1)
Unknown	14 (42.4)
Total	33 (100.0)

classified as inaccurate. Of the reasons for inaccuracy, 15 (45.5%) cases were "to meet review criteria of medical insurance benefits", two (6.1%) cases were "diagnostic error" and one (3.0%) case was "incorrect coding habits". As for the department causing any inaccuracy, the highest was "department of inspection for medical insurance benefit" at 48.5% (Table 5 and 6).

DISCUSSION

When a community-based surveillance system for cardiovascular disease is made, the incidence rate can be calculated easily, but presently reliable data from the community are very difficult to get. This study aimed to ascertain the possibility of using medical insurance data as a valuable source of surveillance for cardiovascular disease by investigating the accuracy of disease codes in medical insurance claims. These claims included 258 inpatients who were admitted to 72 hospitals all over the country during the five years between 1993 and 1997.

Although there were limits to the study regarding the methodological difference of medical recording among medical institutions and the biases of investigators,[4,21] the administrative support by the Ministry of Health and Welfare, a business assistance of Korean Medical Record Association, and a positive aid of the Korean Research Group for Cardiovascular Disease Prevention and Control were very helpful in overcoming the barriers for this study. Therefore, experience in the development of diagnostic criteria for MI, a survey form for MI diagnosis confirmation and the educational materials for the investigation are expected to provide reference data for the developing process of a nationwide surveillance system for cardiovascular disease in future.

Regarding previous studies on the accuracy of disease codes in Korean medical insurance claims, Moon et al reported that the agreement rate was less than 20% when comparing the disease codes. Each medical institution had designated the same disease in medical insurance claims of the nationwide locating Medical Insurance Cooperatives of a certain company for the year 1990.[22] The Ministry of Health and Welfare's survey of cancer patients, which was conducted in 1992, reported that the accuracy of disease codes among the medical insurance claims was 35.7%.[23] Song reported in her "Study for the variations of disease codes in medical insurance claims" that the agreement of disease codes between the claims and medical records was 17.1% for cases with a costly medical bill, and 37.6% for cases with a less expensive medical bill.[24]

In a study on cases where the main or subordinate disease codes in the medical insurance claims had been diagnosed as notifiable acute communicable disease (NACD), the accuracy in 2,339 cases of NACD was 10.1%.[25] Kim et al announced that the discrepancy of diagnosed disease between emergency rooms and discharge was 4.2%. More specifically, the discrepancy was 8.2% in the internal medicine department, 1.5% in the surgery department, and 9.4% in the case of cardiovascular disease. Accordingly, the accuracy of diagnosis in studies within Korea showed a substantial variation by diseases from 10.1% to 95.8%.[26]

Reviewing the foreign literature, the accuracy of DRG diagnosis in the discharge data extracted from 7,050 Medicare cases was 80.5%,[27] and the consistency of diagnosis between the medical insurance claims and the medical records of the hypertension was 74%.[28] The accuracy of CVD was 79% in the study on CVA in Rochester, Minnesota, US,[29] whereas the accuracy of disease in the discharge records of CVA in US was only 28%, of which 84% were hemorrhagic CVA.[30] In addition, Rawson et al informed that the consistency of diagnosed disease between the medical insurance claims and the medical records was 69.3%.[12] In this study, the accuracy of MI disease was 76.0% of all study subjects, and 85.6% among cases possible to assess. The rates of this study tended to be higher than those of studies within Korea,[22-25] but were the same or lower than the rates of foreign studies.[12,27-30]

In studies dealing with the discrepant reasons of disease codes, the difference was due to many factors, such as separation of the process of making medical records and claiming medical insurance, complexity of the claiming process, discrepancy of the time between claiming the medical insurance and making the discharge medical records, physicians' indifference, and abundant disease codes.[24] According to the results from Hsia's study,[31] the coding error was 14.7%, and the rate was 50.7% in excessive repayment. The causes of coding error were doctors' obscure expression for diagnosis (62.9%), medical record administrators' error (9.1%), changing the second diagnosis to the first (26.6%), and others (1.4%).

The major reason of the discrepancy in this study was to meet the review criteria of medical insurance benefits, which mostly occurred in the department of inspection for medical insurance benefits. However, the number of study cases, which were only 33 in this study including 14 missing data, was so limited that it might be unreasonable to interpret the results. Further research with enough samples is necessary to gain more accurate results.

For the above reason, establishing and operating a surveillance system for cardiovascular disease will have to be based primarily on medical records with registered medical record administrators, and the disease codes in medical insurance claims seem to be usable as a supplementary data. Although the subjects of this study were patients who visited medical facilities, it is assumed that everyone suffering from a cardiovascular disease did not visit a hospital. Considering this fact, the operation of a surveillance system for cardiovascular disease should be accompanied by community-based registry. We expect this study to be reflected when drawing up plans (establishment of data collecting system or guiding principle for the education of the informers) to increase the accuracy of diagnosed disease, which is the most important point in operating a surveillance system for cardiovascular disease, and to be helpful in improving medical insurance claims and inspecting medical insurance benefits.

ACKNOWLEDGEMENTS

We would like to thank the Department of Health Promotion at the Ministry of Health and Welfare and members of the Korean Medical Record Association who participated in this study on the development of a model for national cardiovascular disease surveillance.

This study was made possible with the active participation by the Korean Research Group for Cardiovascular Disease Prevention and Control.

In addition to the authors of this paper, members who contributed to this research include: Kee Ho Park, Seungjun Wang, Hwasoon Lee, Dept. of Preventive Medicine, Yonsei University Wonju College of Medicine; Tae-Yong Lee, Un-Je Park, Dept. of Preventive Medicine, College of Medicine, Chungnam National University; Kang-Sook Lee, Hyensook Hong, Dept. of Preventive Medicine, Catholic University Medical College; Duk-Hee Lee, Kwang-wook Koh, Dept. of Preventive Medicine, College of Medicine, Kosin University; Sunhee Lee, Yoomi Chae, Dept. of Preventive Medicine, Medical College of Ewha Womens University; Byung Yeol Chun, Dept. of Preventive Medicine and Public Health, School of Medicine, Kyungpook National University; Jin Sook Suh, Korean Medical Record Association.

REFERENCES

1. National Statistical Office. Annual report on the cause of death statistics (based on vital registration). Seoul, Korea: National Statistical Office; 1998.
2. Suh I, Jee SH, Kim IS. Changing pattern of cardiovascular diseases in Korea. Korean J Epidemiol 1993;15:40-6.
3. National Statistical Office. Annual report on the cause of death statistics (based on vital registration). Seoul, Korea: National Statistical Office; 1996.
4. Park JG. Modern epidemiology. 1st ed. Seoul: Yonsei University Press; 1999.
5. Burke GL, Sprafka JM, Folsom AR, Luepker RV, Norsted SW, Blackburn H. Trends in CHD mortality, morbidity and risk factor levels from 1960 to 1986: The Minnesota Heart Survey. Int J Epidemiol 1989;18 Suppl 1:S73-S81.
6. Ives DG, Fitzpatrick AL, Bild DE, Psaty BM, Kuller LH, Crowley PM, et al. Surveillance and ascertainment of cardiovascular events -The cardiovascular health study-. Ann Epidemiol 1995;5:278-85.
7. Suh I, Chun BY, Park JG, Jeon KH. Development of a model for national cardiovascular disease surveillance. Good Health R & D Project, Ministry of Health & Welfare Republic of Korea (1st Year Report), 1999.
8. Thacker SB, Stroup DF, Parrish RG, Anderson HA. Surveillance in environmental public health: Issues, systems and sources. Am J Public Health 1996;86:633-8.
9. Roos LL, Roos NP, Cageorge SM, Nicol JP. How good

are the data?- reliability of one health ared data bank. Med Care 1982;20:266-76.
10. Fisher ES, Whaley FS, Kryshat WM, Malenka DJ, Fleming C, Baron JA, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. Am J Public Health 1992;82:243-8.
11. Lee GS. Diagnosis coding agreement between medical records and medical claim billing data. (A Master's Thesis) Seoul: The Graduate School Seoul National University; 1995.
12. Rawson NSB, Malcolm E. Validity of the recording of ischemic heart disease and chronic obstructive pulmonary disease in the Saskatchewan health care datafiles. Stat Med 1995;14:2627-43.
13. Shin JY. A comparison study of diagnosis code in both medical records and medical claim bills for same discharge patient - Focused on the inpatient 10 frequent diseases - Annual Bulletin of Seoul Health Junior College 1996;16:79-93.
14. Park JG, Kim KS, Kim CB, Lee TY, Lee KS, Lee DH, et al. A nested case-control study on the risk factor of cardiovascular disease in Korean. Good Health R & D Project, Ministry of Health & Welfare Republic of Korea (1st Year Report), 1998.
15. World Health Organization. Manual of the international statistical classification of diseases, injuries, and causes of death (ninth revision). Geneva: 1977.
16. WHO MONICA Project. Myocardial infarction and coronary deaths in the world health organization MONICA project. Registration procedures, event rates and case-fatality rates in 38 populations from 21 countries in four continents. Circulation 1994;90:583-612.
17. The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: Design and objectives. Am J Epidemiol 1989;129:687-702.
18. Luepker RV, Rastam L, Hannan PJ, Murray DM, Gray C, Baker WL, et al. Community education for cardiovascular disease prevention. Morbidity and mortality results form the Minnesota Heart Health Program. Am J Epidemiol 1996;144:351-62.
19. WHO MONICA Project. Mild myocardial infarction. A classification problem in epidemiologic studies. J Clin Epidemiol 1997;50:3-13.
20. Gillum RF, Fortmann ST, Prineas RJ, Kottke TE. International diagnostic criteria for acute myocardial infarction and acute stroke. Am Heart J 1984;108:150-8.
21. Shinar D, Gross CR, Mohr JP, Caplan LR, Price TR, Wolf PA, et al. Inter-observer variability in the assessment of neurologic history and examination in the stroke data bank. Arch Neurol 1985;42:557-65.
22. Moon OR, Kim CY, Kim KM. Individual variations in the code of the international classification of disease for similar outpatient conditions among general practitioners. Korean J Health Policy Adm 1992;2:66-79.
23. Ministry of Health & Welfare Republic of Korea. 1992 Survey Report on the Cancer Patient, 1996.
24. Song YJ. A study on the diagnostic codes' variation of health insurance claim data. (A Master's Thesis) Seoul: Department of Health Administration, The Graduate School of Yonsei University; 1997.
25. Shin EC, Park YM, Park YG, Kim BS, Park KD, Meng KH. Estimation of disease code accuracy of national medical insurance data and the related factors. Korean J Preventive Med 1998;31:471-80.
26. Kim KH, Seo SW, Won SY, Park SG, Kim SY, Song HS, et al. Survey on discordance rate between final principal diagnosis and principal diagnosis at emergency room. Journal of Korean Society Quality Assurance in Health Care 1998;5:216-23.
27. Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for medicare patients under the prospective-payment system. N Engl J Med 1988;318:352-5.
28. Quam L, Ellis LBM, Venus P, Clouse J, Taylor CG, Leatherman S. Using claims data for epidemiological research: The concordance of claims-based criteria with medical record and patient survey for identifying a hypertensive population. Med Care 1993;31:498-507.
29. Leibson CL, Naessens JM, Brown RD, Whisnant JP. Accuracy of hospital discharge abstracts for identifying stroke. Stroke 1994;25:2348-55.
30. Rosamond WD, Folsom AR, Chambless LE, Wang CH, McGovern PG, Howard G, et al. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the atherosclerosis risk in communities (ARIC) cohort. Stroke 1999;30:736-43.
31. Hsia DC, Ahern CA, Ritchie BP, Moscoe LM, Krushat WM. Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988. JAMA 1992;268: 896-9.