

Evaluation of Factors Influencing Accuracy of Principal Procedure Coding Based on ICD-9-CM: An Iranian Study

by Mehrdad Farzandipour and Abbas Sheikhtaheri

Abstract

To evaluate the accuracy of procedural coding and the factors that influence it, 246 records were randomly selected from four teaching hospitals in Kashan, Iran. “Recodes” were assigned blindly and then compared to the original codes. Furthermore, the coders’ professional behaviors were carefully observed during the coding process. Coding errors were classified as major or minor. The relations between coding accuracy and possible effective factors were analyzed by χ^2 or Fisher exact tests as well as the odds ratio (OR) and the 95 percent confidence interval for the OR.

The results showed that using a tabular index for rechecking codes reduces errors (83 percent vs. 72 percent accuracy). Further, more thorough documentation by the clinician positively affected coding accuracy, though this relation was not significant. Readability of records decreased errors overall ($p = .003$), including major ones ($p = .012$). Moreover, records with no abbreviations had fewer major errors ($p = .021$). In conclusion, not using abbreviations, ensuring more readable documentation, and paying more attention to available information increased coding accuracy and the quality of procedure databases.

Key words: clinical coding, procedures, accuracy, validity, ICD-9-CM, database, Iran

Introduction

Clinical coding quality has become increasingly more important for statistical purposes, reimbursement, administration, epidemiology, and health services research. This research is aimed at evaluating the factors that influence the accuracy of principal procedure coding.

Background

Information must be trustworthy to be used as a valuable resource for the delivery of healthcare.¹ Huffman (1994) argues that classifying information through coding of patients’ diagnoses and procedures will make it more useful.² The International Classification of Diseases (ICD) codes have been used to classify morbidity and mortality information for statistical purposes, administration, epidemiology, and health services research.³ ICD codes are used to track workloads and length of stay, to assess quality of care, to track utilization rates, and to investigate population status and its determinants.^{4,5} Regarding the development of diagnosis-related group (DRG)-based systems in a number of countries such as the United States, Canada, Australia, Germany, Japan, and others, the quality of coding is now known as an important factor in reimbursement.⁶⁻⁹

The increased use of encoded procedural data has resulted in greater attention to code accuracy.^{10, 11} Consequently, accuracy of coded data has been assessed by many investigators, and coding errors have been reported in many research studies.¹²⁻²¹ Accuracy of procedure coding between 53 percent and 100 percent (average 97 percent) was reported in the United Kingdom.²² A study in Saudi Arabia reported 30 percent coding errors, and another study revealed 85–95 percent accuracy in coding of procedures.^{23, 24} Moreover, coding accuracy for cataract surgery in Medicare Part B was 99 percent.²⁵ Additionally, misclassification of procedures can result in over- or underestimation of frequency of procedures.²⁶⁻³⁰

According to the literature, factors such as variance in clinicians' descriptions of procedures, clarity of documentation, incomplete documentation in medical records, use of synonyms and abbreviations to describe the same conditions, lack of physicians' attention to principles of documentation, and coders' experience and education can lead to miscoding.³¹⁻³³ In addition, differences between electronic and paper records, quality assurance programs, indexing errors, lack of coders' attention to ICD principles, and key aspects of the code assignment process have been discussed.³⁴⁻³⁹

There is agreement on the effects of coders' experience and education as well as completeness of documentation on the quality of coding. Moreover, the authors hold the opinion that other factors, such as reviewing the entire medical record and lack of memory-based coding, can improve the quality of coding as well. Many of these factors have been well illustrated in the literature, but there is little research-based knowledge about them.^{40, 41} Furthermore, the discrepancies between coders' perspectives about influential factors and the factors that really affect coding quality necessitate more research on the roots of coding errors with a focus on the coding process.⁴² Because of the lack of sufficient study related to coding accuracy in Iran, scant research data related to factors that affect coding, and the focus of current studies on diagnostic coding, this study aimed to evaluate the accuracy of principal procedure coding and its effective factors in teaching hospitals affiliated with Kashan University of Medical Sciences (KAUMS) in Kashan, Iran.

Research Questions

The study addressed the following questions:

1. To what extent are the errors in principal procedure codes major?
2. What, if any, statistical relationships exist between accuracy of principal procedure coding and the education and experience of coders?
3. Is there any statistical difference between general and specialized hospitals with regard to the accuracy of principal procedure coding?
4. Does better documentation (more thorough information, greater readability, and lack of abbreviations) increase the accuracy of principal procedure coding?
5. Does using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) book increase the accuracy of principal procedure codes as compared with memory-based coding?
6. Does reviewing all information contained in records increase the accuracy of principal procedure coding?

Methods

Study subjects

This study was carried out in all teaching hospitals affiliated with KAUMS, including one general hospital and three specialized ones (psychiatry, ENT, and gynecology). The bed capacities of these hospitals are 418 for the general hospital and 48, 60, and 81, respectively, for the specialized ones. In all of these hospitals, patients' medical records are coded manually after discharge based on ICD-9-CM (vol. 3) in the coding unit of the medical record department. Finally, these codes are stored in a computerized database called the Admission, Discharge System (ADS).

Considering the effect of rare procedures on decreasing coding accuracy, samples were selected with regard to the frequency of procedures in each hospital.⁴³ To this end, we calculated the frequency of principal procedures in each hospital through the ADS based on ICD-9-CM chapters (spring to summer 2007). Finally, 246 medical records were randomly selected in proportion to the bed capacity of the hospitals and procedural frequency per ICD-9-CM chapters from fall 2007 to winter 2008.

Study design

The medical records were reviewed and abstracted in two stages. In the first stage, the principal procedures and their original codes were abstracted in separate checklists. In this stage, the samples were randomly selected from the records forwarded to the coding units and were abstracted immediately after original coding. Original coders and their professional behaviors were observed. At the same time, coders' experience and education, their professional behaviors (reviewing all information contained in records vs. coding based only on face sheet, and using the ICD-9-CM book vs. memory-based coding), necessary information for coding (nature, topography, direction, device and approach of procedures), procedural documentation and use of abbreviations, and readability of records were also recorded in the checklists. "Using the ICD-9-CM book" was defined as coding the records through the alphabetical index of the ICD-9-CM book only or through the alphabetical index accompanied by rechecking of the codes in the tabular list (Figure 1). In this stage, to prevent researcher bias, four well-trained medical record students (each with a bachelor's degree) who were doing their practical training in the coding units abstracted the records. Therefore, the original coders were aware of the abstracting process but were unaware that their behaviors were being observed. Because of ethical considerations, all coders were informed about the study after completing it.

Because of disagreement among coders with regard to abstracting the records, a faculty member of the medical record department in KAUMS (member A) reabstracted all the records blindly.⁴⁴ Due to a number of errors, including changes to the record numbers, misfiling, and misrecording of the numbers in the first-phase abstracts, 60 records were not available for reabstracting. Thus, these abstracts were excluded and subsequently replaced by other records from the same ICD-9-CM chapters based on the above-mentioned process.

Then, a PhD student in health information management (HIM) with five years' experience in clinical coding (member B) compared the abstracts blindly. In this stage, 11 records were replaced due to the major disagreements observed. These disagreements happened because member A abstracted another patient's encounter due to missed admission dates in the first abstracts. Member B compared principles of documentation in the abstracts using the alphabetical index (vol. 3) of ICD-9-CM. Documentation of items such as the nature, topography, direction, device, and approach of procedures was classified as "necessary" or "unnecessary." Documentation of items that appear in ICD-9-CM qualifiers with a specialized code was considered "necessary," and documentation of others was considered "unnecessary." Documentation of items without a qualifier and specialized code was also classified as "unnecessary." Finally, the necessary items were identified as "documented" or "undocumented" based on two abstracts. In case of disagreement between these abstracts, the procedure reabstracted by member A was selected.

Unaware of the original codes, member B recoded all the confirmed abstracts with the same version of ICD-9-CM manually. To calculate reliability, 30 records were selected randomly. Member B recoded these records one month after the first recoding, and no disagreement was found. To calculate interrater reliability, these 30 records were recoded by an assistant professor of HIM who was unaware of the codes and recodes. The reliability between them was 100 percent in the main-digit level, and only three disagreements in decimals of codes were found.

Analytical techniques

The recodes (as gold standards) were compared with the codes identified by the original coders. Therefore, coding accuracy was defined in terms of agreement between original codes and recodes. To calculate coding accuracy, the whole code (the main two digits and two decimals) was considered. Coding errors were classified as "major" or "minor." The major errors were those occurring in the nature or topography of procedures. Coding errors observed in approach, direction, or device were considered

minor. For example, miscoding “cholecystectomy (51.22)” instead of “partial cholecystectomy (51.21)” was regarded as a minor error, and miscoding “reduction of fracture of femur (79.06)” instead of “open reduction of fracture of femur with internal fixation (79.36)” was considered to be a major one.

Because of the subjectivity of abbreviation clarity and record readability, only the abstracts on which the abstractors agreed in terms of abbreviation clarity and readability were considered eligible for abbreviation and readability analysis. The statistical relations were analyzed using SPSS software through χ^2 or Fisher exact tests, the odds ratio (OR), and the 95 percent confidence interval (CI₉₅) for the odds ratio. All analyses were two-sided. Figure 2 shows the entire study protocol.

Results

During the study, eight original coders participated. Among them, seven had a bachelor’s degree, and the other had an associate’s degree in medical records. In addition, five coders were less experienced (five years or less) in medical record departments.

Of the records that were coded, 169 (68.7 percent) were from the general hospital. The less experienced coders coded most of the records (190; 77.2 percent), and coders with a bachelor’s degree coded 225 records (91.5 percent). Coders coded 106 records (43.1 percent) using the ICD-9-CM book and the rest (138; 56.1 percent) in a memory-based manner. They controlled only six codes (2.4 percent) by the ICD-9-CM tabular list. There were 182 records (74.0 percent) coded by reviewing the entire medical record, and the rest (60; 24.4 percent) were coded based on only face sheets.

The procedure nature was documented in all records. After excluding “unnecessary” items, the topography, approach, and direction (or device) were documented in 189 (100 percent), 93 (56.7 percent), and 35 (62.5 percent) of the records, respectively. The abstractors in the two phases agreed on the readability of 223 records (90.7 percent), of which 187 (83.9 percent) were regarded as readable. Additionally, they agreed upon abbreviation clarity of 220 records (89.4 percent). No abbreviations were used to document procedures in 121 records (55.0 percent). Abbreviations in the remaining 99 records (45.0 percent) were clear.

The study showed that 96.7 percent of codes were accurate in the main-digit level, and 81.3 percent of codes were accurate in the four-digit level. Among the 46 errors, 25 (54.3 percent) were major and 21 (45.7 percent) were minor (Table 1). As Table 2 shows, codes assigned by more experienced coders were significantly more accurate than those assigned by less experienced ones (92.9 percent vs. 77.9 percent accuracy; $p = .012$). The coder with an associate’s degree had better codes as well (95.2 percent vs. 80 percent accuracy; $p = .139$). In addition, memory-based coding significantly increased coding accuracy in comparison with coding through the ICD-9-CM book (87.7 percent vs. 72.6 percent accuracy; $p = .003$); however, rechecking the codes in the tabular list increased accuracy in comparison with using only the alphabetical index (83.3 percent vs. 72.0 percent accuracy). Coding accuracy based on the face sheet was equal to coding based on review of the entire record. Furthermore, coding in the specialized hospitals was significantly better than coding in the general hospital (93.5 percent vs. 75.5 percent accuracy; $p = .001$). More thorough documentation of approach, direction, or device improved the rates of coding accuracy nonsignificantly. Clear abbreviations and readability of records significantly improved coding accuracy.

As shown in Table 3, less experienced coders and the coder with only an associate’s degree had nonsignificantly more major errors. Moreover, there were no statistical relations between the types of coding errors and either hospital specialty or more thorough documentation. Using the ICD-9-CM book significantly increased major errors (65.5 percent vs. 35.3 percent; $p = .047$). In addition, coding based on the face sheet nonsignificantly increased major errors (81.8 percent vs. 47.1 percent; $p = .079$). There were fewer major errors in readable records and those records without abbreviations ($p = .012$ and $p = .021$, respectively).

Discussion

The findings showed 46 errors (18.7 percent) in principal procedure codes, 25 (54.3 percent) of which were major. Coding accuracy in Kashan (96.7 percent in main digit and 81.3 percent in the whole code) is better than the 70 percent accuracy in coding diagnoses and procedures that has been identified in Saudi Arabia.⁴⁵ In addition, Javitt et al. showed 99 percent and 96 percent accuracy, respectively, for cataract surgery and its approaches in Medicare Part B.⁴⁶ In an Australian study, Henderson et al. reported 85 percent and 81 percent agreement of the principal procedure coding at the five-digit and seven-digit level, respectively, for 1998–1999 and, likewise, 83 percent and 80 percent for 2000–2001.⁴⁷ In Canada, overall agreement between surgeon billing claims and charts was 95.4 percent for most breast surgeries; additionally, agreement between hospital discharge and chart data was 86.2 percent.⁴⁸ A systematic analysis in the United Kingdom suggests that procedure coding accuracy is 53 percent to 100 percent (average 97 percent).⁴⁹ Moreover, van Walraven and Demers noted that agreements for procedure codes range from 85 to 95 percent.⁵⁰ Because of inconsistencies in code accuracy definition (e.g., number of digits checked), the results of this study cannot be compared to those of other studies. However, it seems that procedure coding accuracy in Kashan is almost equal to that found in other studies. These findings suggest that the databases in Kashan can be trusted; however, more attention is needed to reduce the major errors.

The findings showed that the less experienced coders had more errors; however, these errors were mainly major. According to our definition of the types of errors, it seems that less experienced coders have more misunderstanding about procedures. Therefore, they should pay more attention to the nature and topography of procedures. The coder with an associate's degree had nonsignificantly fewer errors; however, these errors were mainly major. Five of the seven coders with a bachelor's degree who participated in the study were less experienced, so the relationship between coders' education and their coding accuracy may have intervened with coders' experience. In addition, only one coder with an associate's degree, who was more experienced and employed in the psychiatry hospital, participated in the study. The smaller number of procedural codes in the psychiatry hospital as well as his experience may have led to more accuracy. Therefore, this part of the results cannot be generalized to other populations and requires further study.

The findings suggest that using the ICD-9-CM book decreases coding accuracy ($p = .003$), but fewer coding errors happen in those records rechecked with the tabular list as compared to using only the alphabetical index (16.7 percent vs. 28.0 percent error rate). In addition, using the ICD-9-CM book increased major errors ($p = .047$). The effect of using the ICD book on coding accuracy has not been reported, though it has been emphasized strongly in educational materials for coders. We can discuss two possible reasons for this finding. First, some studies have shown that coding accuracy decreases in rare and complicated cases.⁵¹ Therefore, the effect of using the ICD-9-CM book on increasing coding errors is probably due to the coders' use of the ICD-9-CM book only for complicated cases or cases with which they have less experience. Second, the analysis showed that most codes assigned by more experienced coders were significantly based on memory in comparison with those assigned by less experienced coders ($p < .0001$). Since ICD-9-CM has been replaced by ICD-10 for coding diagnoses during the past 10 years, the Iranian educational system seems to have placed little emphasis on teaching procedure coding based on ICD-9-CM.

The accuracy of procedure coding through reviewing the entire medical record was equal to coding based on the face sheet. A study showed that discrepancies in morphology coding were mainly due to inferences about morphology in the absence of microscopic confirmation.⁵² On the other hand, in a study about urological procedures, the authors suggested that one way to reduce coding errors is to make coders use the discharge summary rather than reading through the medical record themselves.⁵³ In addition, coding based on the face sheet increased major errors (81.8 percent vs. 47.1 percent major); therefore, coders are advised to pay more attention to the entire record to understand the procedures well.

Procedure codes in specialized hospitals were significantly more accurate (93.5 percent vs. 75.7 percent accuracy), whereas most errors in the general hospital were major (56.1 percent vs. 40.0 percent major). Other studies have similarly shown differences in coding at various types of hospitals. For example, one study showed that completeness of registration of primary total hip arthroplasty is less in

teaching and low-volume hospitals.⁵⁴ Additionally, teaching and for-profit hospitals report higher percentages of CT and MRI codes.⁵⁵ In contrast to these findings, a study in Australia revealed no difference between general and specialized hospitals.⁵⁶

During the study, five coders with little experience coded records in the general hospital rotationally. It can be concluded that the shorter work periods in the coding unit may result in fewer interactions with physicians and hence little knowledge about procedures, which might lead to major errors. Moreover, the small total of procedures in the specialized hospitals (17 vs. 77 procedure codes) seems to have resulted in improved procedure coding. Pinfold et al. showed that physicians' billing data (codes) for breast surgical procedures are better than hospital discharge abstracts. They concluded that having a limited number of procedures coded by physicians is likely to increase coding accuracy. Furthermore, hospital coders might not know many specifics of the procedures.⁵⁷

In addition, the findings suggest that more thorough documentation improved coding accuracy nonsignificantly. It has been shown that a computerized database, even with better documentation, could not improve the accuracy of procedure coding.^{58, 59} In spite of these studies, several studies suggest that poor documentation may result in decreased code quality.⁶⁰⁻⁶² For example, in one study, the quality of documentation resulted in a 21 percent increase in coding accuracy.⁶³ The nonsignificant improvement in coding accuracy in the current study may be related to the structure of ICD-9-CM. In many records, documentation of approach, device, or direction of procedures was considered unnecessary. This means that there were not any specialized codes for these procedures. Indeed, a study has documented the weaknesses of ICD-9-CM for classifying specialized procedures.⁶⁴

In the current study, records with less readability had significantly more errors (OR = 3.21; CI₉₅: 1.46-7.1), including major ones (OR = 0.14; CI₉₅: 0.02-0.73), in comparison with readable records. A study suggests that the use of typewritten operative notes decreases coding errors that arise from illegible and incomplete handwritten notes.⁶⁵ Clear abbreviations significantly reduced errors ($p = .002$); however, most of these errors were major ($p = .021$). Therefore, it seems that abbreviations are likely to increase misunderstandings about procedures.

Some limitations should be considered in interpreting the findings. In the first stage of the study, the information was concurrently gathered during the original coding period. Although four well-trained medical record students abstracted the records and original coders were unaware that their professional behaviors were being observed, the possibility of researcher bias cannot be neglected. In addition, we could not control for coders' professional behaviors. Therefore, more studies are required to support the findings in terms of these variables. Second, medical records were abstracted in two stages and recoded by a PhD student in HIM with sufficient experience in coding and good reliability. However, the abstracting and recoding processes cannot be considered error-free. Third, the lack of diversity of hospitals affiliated with the university did not allow for generalization of results as related to coders' education and experience or the hospital specialty. Therefore, further studies in more diverse populations will be necessary.

Conclusion

In conclusion, this study suggests that less experienced coders should pay more attention to procedure nature and topography to improve their coding quality. In addition, lack of memory-based coding did not improve coding accuracy. Hence, it is recommended that coders consult with physicians about cases in which the coder has limited knowledge. Moreover, more readable documentation and avoidance of abbreviations by clinicians are recommended.

Mehrdad Farzandipour is an assistant professor of health information management and a member of the faculty of paramedical sciences at Kashan University of Medical Sciences in Kashan, Iran.

Abbas Sheikhtaheri is a PhD candidate in health information management at Iran University of Medical Sciences in Tehran, Iran.

Financial Support

The research was funded by a grant (grant number = 8611) from the vice chancellor for research at Kashan University of Medical Sciences in Kashan, Iran.

Acknowledgments

The authors thank S. Yarahmadi, N. Sayadi, Z. Nasaj, M. Keshavarz, and L. Shokrizade, MS, for their great help in data gathering. In addition, the authors thank all the coders who participated in the study.

Notes

1. Lorence, D. P. "Regional Variation in Medical Classification Agreement: Benchmarking the Coding Gap." *Journal of Medical Systems* 27, no. 5 (2003): 435–443.
2. Huffman, E. *Health Information Management*. 10th ed. Berwyn, IL: Physician's Record Company, 1994.
3. O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C.M. Ashton. "Measuring Diagnoses: ICD Code Accuracy." *Health Services Research* 40, no. 5, pt. 2 (2005): 1620–1639.
4. Quan, H., G. A. Pearsons, and W. A. Ghali. "Validity of Procedure Codes in International Classification of Diseases, 9th Revision, Clinical Modification Administrative Data." *Medical Care* 42, no. 8 (2004): 801–809.
5. O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."
6. Canadian Institute for Health Information. Acute Care Grouping Methodologies: From Diagnosis Related Groups to Case Mix Groups Redevelopment. 2004. Available at http://secure.cihi.ca/cihiweb/products/Acute_Care_Grouping_Methodologies2004_e.pdf (retrieved February 15, 2009).
7. Australian Department of Health and Aging. Australian Refined Diagnosis Related Groups (AR-DRGs). 2008. Available at www.health.gov.au:80/internet/main/Publishing.nsf/Content/Classificationshub.htm (retrieved February 15, 2009).
8. Hitoshi, K. "Japan Edition Diagnosis Related Group." *Modern Medical Laboratory* 28, no. 11 (2000): 1375–1378.
9. Aileen, R. N., et al. "A New Concept for DRG-based Reimbursement of Services in German Intensive Care Units: Results of a Pilot Study." *Intensive Care Medicine* 30, no. 6 (2004): 1220–1223.
10. Lorence, D. P. "Regional Variation in Medical Classification Agreement: Benchmarking the Coding Gap."
11. O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."
12. Li, R., Z. Hao, and K. X. Liu. "Analysis of the Quality of Coding for Tumors in Line with International Classification of Diseases." *Di Yi Jun Yi Da Xue Xue Bao* 24, no. 2 (2004): 187–191.
13. Hasan, M., R. J. Meara, and B. K. Bhowmick. "The Quality of Diagnostic Coding in Cerebrovascular Disease." *International Journal for Quality in Health Care* 7, no. 4 (1995): 407–410.
14. Bazarian, J. J., P. Veazie, S. Mookerjee, and E. B. Lerner. "Accuracy of Mild Traumatic Brain Injury Case Ascertainment Using ICD-9." *Academic Emergency Medicine* 13, no. 1 (2006): 31–38.
15. MacIntyre, C. R., M. J. Ackland, E. J. Chandraraj, and J. E. Pilla. "Accuracy of ICD-9-CM Codes in Hospital Morbidity Data, Victoria: Implications for Public Health Research." *Australian and New Zealand Journal of Public Health* 21, no. 5 (1997): 477–482.
16. Faciszewski, T., R. Jensen, and R. L. Berg. "Procedural Coding of Spinal Surgeries (CPT-4 Versus ICD-9-CM) and Decisions Regarding Standards: A Multicenter Study." *Spine* 28, no. 5 (2003): 502–507.
17. Stausberg, J., N. Lehmann, D. Kaczmarek, and M. Stein. "Reliability of Diagnoses Coding with ICD-10." *International Journal of Medical Informatics* 77, no. 1 (2008): 50–57.

18. Kern, E. F., M. Maney, D. R. Miller, C. L. Tseng, A. Tiwari, M. Rajan, D. Aron, and L. Pogach. "Failure of ICD-9-CM Codes to Identify Patients with Comorbid Chronic Kidney Disease in Diabetes." *Health Services Research* 41, no. 2 (2006): 564–580.
19. Gance, L. G., A. W. Dick, T. M. Osler, and D. B. Mukamel. "Does Date Stamping ICD-9-CM Codes Increase the Value of Clinical Information in Administrative Data?" *Health Services Research* 41, no. 1 (2006): 231–251.
20. Quan, H., B. Li, L. D. Saunders, G. A. Parsons, C. I. Nilsson, A. Alibhai, and W. A. Ghali. "Assessing Validity of ICD-9-CM and ICD-10 Administrative Data in Recording Clinical Conditions in a Unique Dually Coded Database." *Health Services Research* 43, no. 4 (2008): 1424–1441.
21. Rangachari, P. "Coding for Quality Measurement: The Relationship between Hospital Structural Characteristics and Coding Accuracy from the Perspective of Quality Measurement." *Perspectives in Health Information Management* 4, no. 3 (2007). Available at www.ahima.org/perspectives (retrieved November 20, 2008).
22. Campbell, S. E., M. K. Campbell, J. M. Grimshaw, and A. E. Walker. "A Systematic Review of Discharge Coding Accuracy." *Journal of Public Health Medicine* 23, no. 3 (2001): 205–211.
23. Farhan, J., S. Al-Jummaa, A. A. Alrajhi, H. Al-Rayes, and A. Al-Nasser. "Documentation and Coding of Medical Records in a Tertiary Care Center: A Pilot Study." *Annals of Saudi Medicine* 25, no. 1 (2005): 46–49.
24. Van Walraven, C., and S. V. Demers. "Coding Diagnoses and Procedures Using a High-Quality Clinical Database Instead of a Medical Record Review." *Journal of Evaluation in Clinical Practice* 7, no. 3 (2001): 289–297.
25. Javitt, J. C., A. M. McBean, S. S. Sastry, and F. DiPaolo. "Accuracy of Coding in Medicare Part B Claims: Cataract as a Case Study." *Archives of Ophthalmology* 111, no. 5 (1993): 605–607.
26. Quan, H., G. A. Pearsons, and W. A. Ghali. "Validity of Procedure Codes in International Classification of Diseases, 9th Revision, Clinical Modification Administrative Data."
27. De Coster, C., B. Li, and H. Quan. "Comparison and Validity of Procedures Coded with ICD-9-CM and ICD-10-CA/CCI." *Medical Care* 46, no. 6 (2008): 627–634.
28. Dismuke, C. E. "Underreporting of Computed Tomography and Magnetic Resonance Imaging Procedures in Inpatient Claims Data." *Medical Care* 43, no. 7 (2005): 713–717.
29. Qureshi, A. I., P. Harris-Lane, F. Siddiqi, and J. F. Kirmani. "International Classification of Diseases and Current Procedural Terminology Codes Underestimated Thrombolytic Use for Ischemic Stroke." *Journal of Clinical Epidemiology* 59 (2006): 856–858.
30. Henderson, T., J. Shephard, and V. Sundararajan. "Quality of Diagnosis and Procedure Coding in ICD-10 Administrative Data." *Medical Care* 44, no. 11 (2006): 1011–1019.
31. O'Malley, K. J., K. F. Cook, M.D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."
32. Quan, H., G. A. Pearsons, and W. A. Ghali. "Validity of Procedure Codes in International Classification of Diseases, 9th Revision, Clinical Modification Administrative Data."
33. Hasan, M., R. J. Meara, and B. K. Bhowmick. "The Quality of Diagnostic Coding in Cerebrovascular Disease."
34. Van Walraven, C., and S. V. Demers. "Coding Diagnoses and Procedures Using a High-Quality Clinical Database Instead of a Medical Record Review."
35. Silfen, E. "Documentation and Coding of ED Patient Encounters: An Evaluation of the Accuracy of an Electronic Medical Record." *American Journal of Emergency Medicine* 24, no. 6 (2006): 664–678.
36. O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."

37. Hasan, M., R. J. Meara, and B. K. Bhowmick. "The Quality of Diagnostic Coding in Cerebrovascular Disease."
38. Surján, G. "Questions on Validity of International Classification of Diseases-Coded Diagnoses." *International Journal of Medical Informatics* 54, no. 2 (1999): 77–95.
39. Abdelhak, M. (Editor). *Health Information: Management of a Strategic Resource*. 2nd ed. Philadelphia: W. B. Saunders, 2001.
40. O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."
41. Surján, G. "Questions on Validity of International Classification of Diseases-Coded Diagnoses."
42. Santos, S., G. Murphy, K. Baxter, and K. M. Robinson. "Organisational Factors Affecting the Quality of Hospital Clinical Coding." *Health Information Management Journal* 37, no. 1 (2008): 25–37.
43. Campbell, S. E., M. K. Campbell, J. M. Grimshaw, and A. E. Walker. "A Systematic Review of Discharge Coding Accuracy."
44. O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."
45. Farhan, J., S. Al-Jummaa, A. A. Alrajhi, H. Al-Rayes, and A. Al-Nasser. "Documentation and Coding of Medical Records in a Tertiary Care Center: A Pilot Study."
46. Javitt, J. C., A. M. McBean, S. S. Sastry, and F. DiPaolo. "Accuracy of Coding in Medicare Part B Claims: Cataract as a Case Study."
47. Henderson, T., J. Shepherd, and V. Sundararajan. "Quality of Diagnosis and Procedure Coding in ICD-10 Administrative Data."
48. Pinfold, S. P., V. Goel, and C. Sawka. "Quality of Hospital Discharge and Physician Data for Type of Breast Cancer Surgery." *Medical Care* 38, no. 1 (2000): 99–107.
49. Campbell, S. E., M. K. Campbell, J. M. Grimshaw, and A. E. Walker. "A Systematic Review of Discharge Coding Accuracy."
50. Van Walraven, C., and S. V. Demers. "Coding Diagnoses and Procedures Using a High-Quality Clinical Database Instead of a Medical Record Review."
51. Campbell, S. E., M. K. Campbell, J. M. Grimshaw, and A. E. Walker. "A Systematic Review of Discharge Coding Accuracy."
52. Brewster, D., C. Muir, and J. Crichton. "Registration of Non-Melanoma Skin Cancers in Scotland—How Accurate Are Site and Morphology Codes?" *Clinical and Experimental Dermatology* 20, no. 5 (1995): 401–405.
53. Khwaja, H. A., H. Syed, and D. W. Cranston. "Coding Errors: A Comparative Analysis of Hospital and Prospectively Collected Departmental Data." *BJU International* 89, no. 3 (2002): 178–180.
54. Pedersen, A., S. Johnsen, S. Overgaard, K. Soballe, H. T. Sorensen, and U. Lucht. "Registration in the Danish Hip Arthroplasty Registry: Completeness of Total Hip Arthroplasties and Positive Predictive Value of Registered Diagnosis and Postoperative Complications." *Acta Orthopaedica Scandinavica* 75, no. 4 (2004): 434–441.
55. Dismuke, C. E. "Underreporting of Computed Tomography and Magnetic Resonance Imaging Procedures in Inpatient Claims Data."
56. Santos, S., G. Murphy, K. Baxter, and K. M. Robinson. "Organisational Factors Affecting the Quality of Hospital Clinical Coding."
57. Pinfold, S. P., V. Goel, and C. Sawka. "Quality of Hospital Discharge and Physician Data for Type of Breast Cancer Surgery."

58. Van Walraven, C., and S. V. Demers. "Coding Diagnoses and Procedures Using a High-Quality Clinical Database Instead of a Medical Record Review."
59. Silfen, E. "Documentation and Coding of ED Patient Encounters: An Evaluation of the Accuracy of an Electronic Medical Record."
60. Hasan, M., R. J. Meara, and B. K. Bhowmick. "The Quality of Diagnostic Coding in Cerebrovascular Disease."
61. Geller, S. E., S. Ahmed, M. L. Brown, S. M. Cox, D. Rosenberg, and S. J. Kilpatrick. "International Classification of Diseases 9th Revision Coding for Preeclampsia: How Accurate Is It?" *American Journal of Obstetrics and Gynecology* 190, no. 6 (2004): 1629–1633.
62. McKenzie, K., S. Walker, L. C. Dixon, G. Dear, and F. J. Moran. "Clinical Coding Internationally: A Comparison of the Coding Workforce in Australia, America, Canada, and England." IFHRO Congress & AHIMA Convention Proceedings. 2004. Available at www.ahima.org (retrieved November 21, 2007).
63. Farhan, J., S. Al-Jummaa, A. A. Alrajhi, H. Al-Rayes, and A. Al-Nasser. "Documentation and Coding of Medical Records in a Tertiary Care Center: A Pilot Study."
64. Faciszewski, T., R. Jensen, and R. L. Berg. "Procedural Coding of Spinal Surgeries (CPT-4 Versus ICD-9-CM) and Decisions Regarding Standards: A Multicenter Study."
65. Arthur, J., and R. Nair. "Increasing the Accuracy of Operative Coding." *Annals of the Royal College of Surgeons of England* 86, no. 3 (2004): 210–212.

Figure 1

Checklist Used for Gathering Information

| Abstract form | | | |
|---|--|--|---|
| Hospital: | Record number: | Admission date: | |
| Procedure done: | | | |
| Nature | Not documented <input type="checkbox"/> | Topography: | Not documented <input type="checkbox"/> |
| Approach: | Not documented <input type="checkbox"/> | Device: | Not documented <input type="checkbox"/> |
| Direction: | One-sided <input type="checkbox"/> | Two-sided <input type="checkbox"/> | Not documented <input type="checkbox"/> |
| Other information: | | | |
| Readability: | Readable <input type="checkbox"/> | Unreadable <input type="checkbox"/> | |
| Abbreviations: | Clear <input type="checkbox"/> | Obscure <input type="checkbox"/> | Without abbreviation <input type="checkbox"/> |
| Studying the information: | Only face sheet <input type="checkbox"/> | Face sheet and other forms <input type="checkbox"/> | |
| Did the coder code this record with the ICD-9-CM book? | | | |
| YES: | Only alphabetical index <input type="checkbox"/> | Alphabetical index and rechecking in tabular list <input type="checkbox"/> | |
| No: | (Memory-based coding) <input type="checkbox"/> | | |
| Coder's education: | Associate's degree <input type="checkbox"/> | Bachelor's degree <input type="checkbox"/> | |
| Coder's experience: | ≤ five years <input type="checkbox"/> | > five years <input type="checkbox"/> | |

Figure 2

Entire Study Protocol

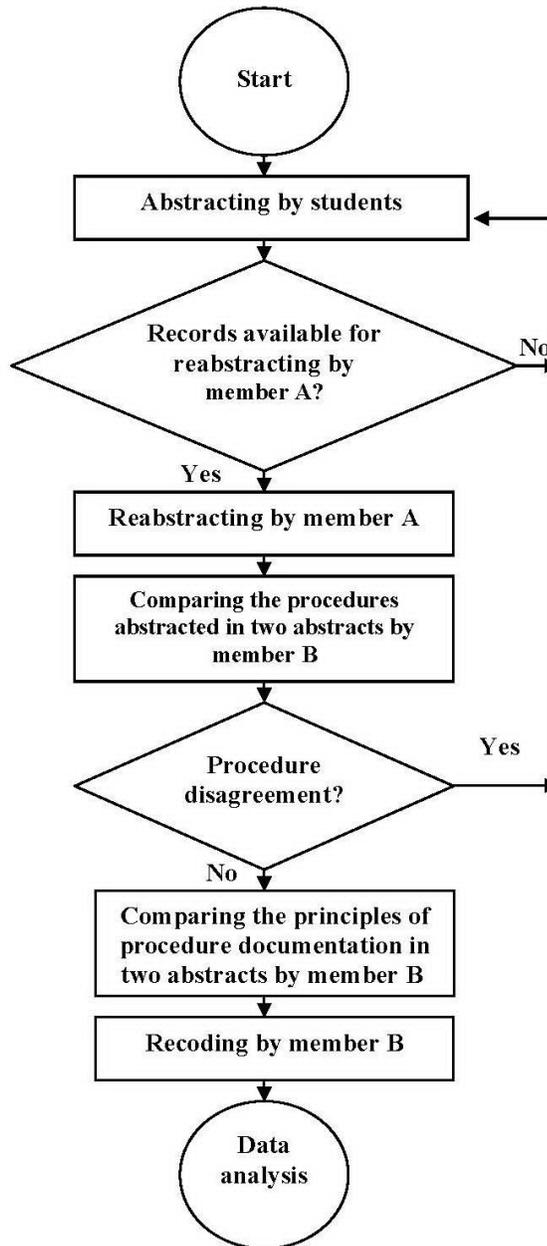


Table 1

Principal Procedure Code Accuracy in Teaching Hospitals Affiliated with Kashan University of Medical Sciences

| | Main digits, frequency (percent) | Decimal digits, frequency (percent) |
|--------------|---|--|
| Accurate | 238 (96.7) | 200 (81.3) |
| Errors | 8 (3.3) | 46 (18.7) |
| Major | 8 (100) | 25 (54.3) |
| Minor | – | 21 (45.7) |
| Total | 246 (100) | 246 (100) |

Note: Percentages of major and minor errors were calculated based on total number of errors.

Table 2

Factors Related to Principal Procedure Code Accuracy

| | | Frequency (percent) | | | χ^2 | <i>p</i> | OR | CI ₉₅ for OR |
|--------------------------|-----------------------------|---------------------|-----------|-------|----------|----------|------|-------------------------|
| | | Accurate | Error | Total | | | | |
| Experience (years)* | ≤5 | 148 (77.9) | 42 (22.1) | 190 | 6.369 | .012 | 0.27 | 0.09–0.77 |
| | >5 | 52 (92.9) | 4 (7.1) | 56 | | | | |
| Education | Bachelor | 180 (80.0) | 45 (20.0) | 225 | 2.934 | .139 | 0.2 | 0.02–1.54 |
| | Associate | 20 (95.2) | 1 (4.8) | 21 | | | | |
| Using the ICD-9-CM book* | Alphabetical only | 72 (72.0) | 28 (28.0) | 100 | 9.339 | .009 | – | – |
| | Alphabetical & tabular list | 5 (83.3) | 1 (16.7) | 6 | | | | |
| | Memory-based | 121 (87.7) | 17 (12.3) | 138 | | | | |
| Using the ICD-9-CM book* | With book | 77 (72.6) | 29 (27.4) | 106 | 8.864 | .003 | 0.37 | 0.19–0.73 |
| | Memory-based | 121 (87.7) | 17 (12.3) | 138 | | | | |
| Record study | Entire review | 148 (81.3) | 34 (18.7) | 182 | 0.004 | .952 | 0.98 | 0.46–2.07 |
| | Face sheet | 49 (81.7) | 11 (18.3) | 60 | | | | |
| Hospital* | General | 128 (75.7) | 41 (24.3) | 169 | 10.984 | .001 | 0.22 | 0.08–0.57 |
| | Specialized | 72 (93.5) | 5 (6.5) | 77 | | | | |
| Nature | Documented | 200 (81.3) | 46 (18.7) | 246 | – | – | – | – |
| | Undocumented | – | – | – | | | | |
| Topography | Documented | 145 (76.7) | 44 (23.3) | 189 | – | – | – | – |
| | Undocumented | – | – | – | | | | |
| Approach | Documented | 71 (76.3) | 22 (23.7) | 93 | 0.207 | .649 | 1.18 | 0.58–2.39 |
| | Undocumented | 52 (73.2) | 19 (26.8) | 71 | | | | |
| Direction or device | Documented | 22 (62.9) | 13 (37.1) | 35 | 0.180 | .672 | 1.27 | 0.42–3.82 |
| | Undocumented | 12 (57.1) | 9 (42.9) | 21 | | | | |
| Readability* | Readable | 159 (85.0) | 28 (15.0) | 187 | 8.989 | .003 | 3.21 | 1.46–7.1 |
| | Unreadable | 23 (63.9) | 13 (36.1) | 36 | | | | |
| Abbreviation* | No abbreviation | 90 (74.4) | 31 (25.6) | 121 | 10 | .002 | 0.29 | 0.13–0.64 |
| | Clear abbreviation | 90 (90.9) | 9 (9.1) | 99 | | | | |

* Significant statistical relation at 95% level

Table 3

Factors Related to Principal Procedure Code Accuracy by Type of Error

| | | Frequency (percent) | | | χ^2 | <i>p</i> | OR | CI ₉₅ for OR |
|---------------------------|-----------------------------|---------------------|-----------|-------|----------|----------|------|-------------------------|
| | | Major | Minor | Total | | | | |
| Experience (years) | ≤5 | 24 (57.1) | 18 (42.9) | 42 | 1.521 | .318 | 4 | 0.38–41.85 |
| | >5 | 1 (25.0) | 3 (75.0) | 4 | | | | |
| Education | Bachelor | 24 (53.3) | 21 (46.8) | 45 | 0.895 | .534 | – | – |
| | Associate | 1 (100) | – | 1 | | | | |
| Using the ICD-9-CM book | Alphabetical only | 19 (67.9) | 9 (32.1) | 28 | 5.738 | .057 | – | – |
| | Alphabetical & tabular list | – | 1 (100) | 1 | | | | |
| | Memory-based | 6 (35.3) | 11 (64.7) | 17 | | | | |
| Using the ICD-9-CM book * | With book | 19 (65.5) | 10 (34.5) | 29 | 3.946 | .047 | 3.48 | 0.99–12.2 |
| | Memory-based | 6 (35.3) | 11 (64.7) | 17 | | | | |
| Record study | Entire review | 16 (47.1) | 18 (52.9) | 34 | 4.067 | .079 | 0.2 | 0.04–1.06 |
| | Face sheet | 9 (81.8) | 2 (18.2) | 11 | | | | |
| Hospital | General | 23 (56.1) | 18 (43.9) | 41 | 0.465 | .648 | 1.92 | 0.29–12.8 |
| | Specialized | 2 (40.0) | 3 (60.0) | 5 | | | | |
| Nature | Documented | 25 (54.3) | 21 (45.7) | 46 | – | – | – | – |
| | Undocumented | – | – | – | | | | |
| Topography | Documented | 24 (54.5) | 20 (45.5) | 44 | – | – | – | – |
| | Undocumented | – | – | – | | | | |
| Approach | Documented | 13 (59.1) | 9 (40.9) | 22 | 1.177 | .278 | 1.99 | 0.32–12.3 |
| | Undocumented | 8 (42.1) | 11 (57.9) | 19 | | | | |
| Direction or device | Documented | 9 (69.2) | 4 (30.8) | 13 | 0.016 | .628 | 1.12 | 0.18–6.96 |
| | Undocumented | 6 (66.7) | 3 (33.3) | 9 | | | | |
| Readability* | Readable | 12 (42.9) | 16 (57.1) | 28 | 6.286 | .012 | 0.14 | 0.02–0.73 |
| | Unreadable | 11 (84.6) | 2 (15.4) | 13 | | | | |
| Abbreviation* | No abbreviation | 13 (41.9) | 18 (58.1) | 31 | 6.166 | .021 | 0.09 | 0.01–0.82 |
| | Clear abbreviation | 8 (88.9) | 1 (11.1) | 9 | | | | |

*Significant statistical relation at 95% level