Window of Health
Jurnal Kesehatan
journal homepage : www.jurnal.fkmumi.ac.id

ARTICLE RESEARCH
URL artikel: http://jurnal.fkmumi.ac.id/index.php/woh/article/view/woh8211

Analysis of the Factors Influencing the Coding Quality in East Java Hospital of Indonesia: Diabetes Mellitus as A Case Study

Hosizah Markam[1], Kori Puspita Ningsih[2], Lily Widjaja [3], [C] Umi Khoirun Nisak [4], Roy Rillera Marzo[5], Siswati[6]

[1]Health Information Management, Health Science Faculty, Universitas Esa Unggul, Indonesia
[2]Medical Record and Health Information, Health Science Faculty, Universitas Jendral Achmad Yani Yogyakarta, Indonesia
[3]Medical Record and Health Information, Health Science Faculty, Universitas Esa Unggul, Indonesia
[4]Health Information Management, Health Science Faculty, Universitas Muhammadiyah Sidoarjo, Indonesia
[5]Faculty of Humanities and Health Sciences, Curtin University Malaysia, Malaysia
[6]Health information management, health science faculty, Universitas Esa Unggul, Indonesia
Email Corresponding Author([C]): hozisah@esaunggul.ac.id.
hozisah@esaunggul.ac.id[1] , puspitakori@gmail.com[2] , lily.widjaja@esaunggul.ac.id[3]
umikhoirun@umsida.ac.id[4], Roy.Marzo@curtin.edu.au[5], siswati@esaunggul.ac.id[6]

ABSTRACT

Code precision has received significant attention due to the increased utilization of encoded procedural data. Coding errors have been documented in multiple research investigations. This study aims to assess the variables that affect coding quality. The prevalence of diabetes has increased substantially in the past two decades and is a significant cause of morbidity and mortality. Method: This study was conducted in 2 hospitals in East Java, Indonesia, that were selected through simple random sampling from a population of hospitals meeting the predefined inclusion criteria. The bed capacities of these hospitals are 211, with details of 62 and 149, respectively, for the specialised ones. The sample in this study was 60 medical record files taken randomly in 2022 in the case of diabetes mellitus. The result showed coding quality testing uses six elements: reliability, accuracy, relevancy, timeliness, completeness, and legibility. Data analysis was carried out analytically using the Fisher Exact test. The results of the study from 60 samples showed that four elements were significant out of a total of 6 aspects of coding quality elements. The four essential elements consisted of Accuracy ($p=0.001$), Reliability ($p=0.001$), Completeness ($p=0.046$), and Legibility ($p=0.046$). Reliability elements also impact coding accuracy or vice versa ($p=0.001$); Completeness also affects Legibility and vice versa ($p=0.046$). The odds ratio value of each component shows that Reliability and Accuracy are 8.782, which means that Reliability can increase Accuracy 8 times and vice versa. Meanwhile, completeness and legibility are at 3.818, which means completeness also increases legibility by three times and vice versa. The Hospitals should consider four significant coding quality elements, including completeness, accuracy, reliability, and legibility, for use in coding audits. Timeliness and Relevance were insignificant.

Keywords : Coding; audit; healthcare; diabetes mellitus

PUBLISHED BY :
Faculty of Public Health
Universitas Muslim Indonesia
Address :
Jl. Urip Sumohardjo Km. 5 (Campus II UMI)
Makassar, Sulawesi Selatan.
Email :
jurnalwoh.fkm@umi.ac.id
Phone :
+62 82188474722

Article history
Received 7 September 2024
Received in revised form 23 April 2025
Accepted 24 April 2025
Available online 26 April 2025

INTRODUCTION

The prevalence of diabetes mellitus has increased substantially in the past two decades and is a significant cause of morbidity and mortality [1]. As the leading cause of blindness, end-stage renal disease, and cardiovascular disease, diabetes poses a significant challenge to the healthcare system. It places a considerable burden on patients and their families [2]. Therefore, diabetes has been studied extensively to project the incidence in the population, identify high-risk groups, and evaluate prevention and control initiatives to reduce the disease and its complications [3–5]. Consequently, diabetes has been extensively studied to project its future incidence, identify high-risk populations, and evaluate the effectiveness of prevention and control initiatives designed to reduce the disease's prevalence and the severity of its complications. Despite the considerable focus on clinical management and prevention, less attention has been paid to the accuracy of disease documentation and coding, which are critical for epidemiological surveillance, resource allocation, and health system planning. Errors in medical coding can lead to underreporting or misclassifying diabetes cases, subsequently affecting the validity of research outcomes and policy decisions. A recent audit of medical coding practices revealed that 18% of diabetes-related cases at RS Siti Fatimah Tulangan and 22% at RS Bhayangkara were incorrectly coded. These inaccuracies stemmed primarily from diagnostic documentation inconsistencies and misinterpretations of ICD-10 classification guidelines. The relatively high miscoding rates observed in these hospitals underscore a systemic gap in clinical documentation and coding practices, suggesting a pressing need for continuous professional development programs, stricter quality control measures, and the integration of automated coding assistance tools. Addressing these challenges is vital to ensure the reliability of diabetes data, optimize healthcare planning, and ultimately improve patient outcomes.

For information to be helpful in healthcare provision, it must be reliable. Huffman (1994) states that data can be more effectively categorized by coding patient diagnoses and procedures [3]. Information on morbidity and mortality has been classified using the International Classification of Diseases (ICD) codes for statistical, administrative, epidemiological, and health services research purposes [6–9] [10]. Three ICD codes are utilized to monitor utilization rates, workloads, length of stay, quality of care evaluation, and population status and its factors [6]. Coding quality is currently recognized as a significant issue in reimbursement regarding the INA-CBGs (Indonesia Case-Based *Groups).*

Code precision has received significant attention due to the increased utilization of encoded procedural data [11]. As a result, numerous investigators have evaluated the precision of coded data, and multiple research investigations have documented coding errors [4]. Reports in the United Kingdom indicated an accuracy of procedure coding ranging from 53% to 100%, with an average of 97%. The accuracy of procedure coding was between 85 and 95% in one study conducted in Saudi Arabia, while another investigation documented coding mistakes of only 30%. Furthermore, Medicare Part B codes cataract surgery with a 99 percent accuracy rate [12].

Furthermore, the underestimating or overestimating procedure frequency may result from incorrect classification. Several studies state the factors that affect the quality of clinical coding, namely concept orientation, consistency, soundness, and non-redundancy[13,14]. The results of preliminary studies obtained at Siti Fatimah Tulangan Hospital were based on preliminary data from 10 medical records of diabetes mellitus cases. Researchers audited the coding of inpatients in private hospitals and government-owned hospitals. The results obtained at the Bhayangkara Pusdik Porong Hospital are based on the reliability variable of the consistency of the results of each clinical coder, as many as 7 (70%). Variable completeness of supporting diagnoses as much as 6 (60%). Variable timeliness of coding 2x24 hours as much as 8 (80%). Variable accuracy of coding accuracy as much as 2 (20%). Variable definition of the suitability of standard abbreviations as much as 9 (90%). The results obtained at the Siti Fatimah Tulangan 'Aisyiyah Hospital were based on the reliability variable of the consistency of the results of each clinical coder, which was 7 (70%). The completeness variable of supporting diagnoses was 5 (50%). Timeliness variable timeliness of coding 2x24 hours as many as 8 (80%). The accuracy variable of coding accuracy is 4 (40%). There are various definitions of the suitability of standard abbreviations, as many as 10 (100%).

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 [15–18]. In addition, differences between electronic and paper records, quality assurance programs, indexing errors, lack of coders' attention to ICD principles, and critical aspects of the code assignment process have been discussed. In previous research, these factors have become coding quality attributes: reliability, completeness, timeliness, accuracy, relevancy, definition, and legibility. Accuracy indicates that data should be correct, right, and consistent. Completeness refers to the point that data should be present and comprehensive. Relevancy, as another attribute, is related to the usability and usefulness of data and the data's fitness for the purpose. Timeliness indicates that data should be timely and current. Definition presents that data should be valid, precise, and understandable and have a clear and unique meaning. Data representation format, by definition, is the format by which data are presented to the end user. In other words, this attribute indicates the body or corpus of data. [19,20]

There is general agreement on the effects of coders' experience, education level, and the completeness of clinical documentation on coding quality. Moreover, the authors believe that additional factors, such as systematic review of medical records and the avoidance of memory-based coding practices, can also contribute significantly to improving coding quality. Many of these factors have been well illustrated in previous literature; however, there remains a need for more research-based knowledge, particularly in diverse healthcare settings. Furthermore, much of the existing research has concentrated predominantly on diagnostic coding, often overlooking procedure coding and specialty-specific coding challenges.

Errors in clinical coding are not trivial; they can lead to substantial problems across multiple domains of healthcare delivery. Misclassification of diagnoses may result in incorrect patient treatment pathways, inappropriate billing and reimbursement, distorted health statistics, flawed clinical audits, and inaccurate epidemiological surveillance. These errors can compromise patient safety, undermine hospital accreditation and reporting systems, and ultimately affect health policy decisions and resource allocations. Erroneous coding can also lead to financial losses for healthcare institutions and distort national health data used for research and planning. Therefore, addressing the factors that influence coding quality is critical to maintaining the integrity of health information systems. This study aims to analyze the determinants of coding quality using the ICD-10 system within selected hospitals in East Java, Indonesia, with Diabetes Mellitus chosen as the case study to represent a high-burden chronic disease frequently encountered in clinical practice.

METHOD

This study utilized an analytic observational cross-sectional design to assess the association between various factors and the quality of medical coding in hospital settings. This study was conducted in 2 hospitals in East Java, Indonesia. The inclusion criteria for the medical coding audit were as follows: Hospitals must be fully operational during the audit period and possess a standardized medical records unit that complies with national standards, such as the Ministry of Health Regulation No. 24 of 2022 on Medical Records. Eligible hospitals must also utilize internationally recognized diagnostic and procedural coding systems, specifically ICD-10 and ICD-9-CM/ICD-10-PCS. Furthermore, hospitals must provide access to electronic or physical inpatient and/or outpatient medical records. Only hospitals accredited at a minimum of the "Madya" (Intermediate) level by the Hospital Accreditation Commission (KARS) or an equivalent accreditation body were included. Hospitals must express their willingness to allow auditing of their medical records while ensuring patient confidentiality by regulations.

Additionally, participating hospitals needed to have recorded at least 100 inpatient or outpatient cases within the past three months to ensure sufficient volume for reliable coding analysis. Hospitals were excluded from the medical coding audit if they were newly established, had not been operational for at least one year, lacked a formalized medical records unit, or had not implemented standard coding systems such as ICD-10 or ICD-9-CM/ICD-10-PCS. Facilities that maintained only manual, unstructured records without systematic coding practices were also excluded. Furthermore, hospitals that declined to grant access to their medical records for auditing purposes or could not ensure patient data's confidentiality and security as mandated by applicable laws and regulations were disqualified. Hospitals with fewer than 100 recorded inpatient or outpatient cases in the preceding three months were similarly excluded to ensure an adequate sample size for valid coding analysis. To ensure impartiality, hospitals were randomly allocated, and two were subsequently selected for the audit process.

The bed capacities of these hospitals are 211, with details of 62 and 149, respectively, for the specialized ones. In all of these hospitals, patients' medical records are coded manually after discharge based on ICD-10 in the coding unit of the medical record department. The sample in this study was 60 medical record files taken randomly in 2022 in the case of diabetes mellitus. The medical records were reviewed and abstracted in two stages. The principal procedures and their original codes were abstracted in separate checklists in the first stage. In this stage, the samples were randomly selected from the records forwarded to the coding units and were abstracted immediately after the original coding. Original coders and their professional behaviors were observed. In this stage, one coder professional coding units abstracted the records to prevent researcher bias. 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 the subjectivity of abbreviation clarity and record readability, only the abstracts on which the abstractors agreed regarding abbreviation clarity and readability were considered eligible for abbreviation and readability analysis. The statistical relations were analyzed using SPSS software through $\chi 2$ or Fisher exact tests, the odds ratio (OR), and the 95 percent confidence interval (CI 95) for the odds ratio. All analyses were two-sided.

RESULTS

The study showed that 53.3% percent of codes were reliable, 70% complete, 65% timelines, 43.3 % accurate, 86.7% relevant, 80% definition, and 78.3% legibility. The percentage shows that most coders are by variables with a rate of more than 50%, such as reliability, completeness, timeliness, relevance, definition, and legibilty. Completeness, reliability, accuracy, and definition influence the quality of clinical coding. [4,21,22] [12]. Of all the variables, it can be seen that the definition variable impacts the other variables.

Table 1. Reliability Variable

Variable Reliability	Freq	Percentage
Reliable	32	53.3
Not Reliable	28	46.7
Total	60	100

Based on the provided in Table 1, Out of 60 cases, 53.3% (32 cases) were deemed reliable, while 46.7% (28 cases) were classified as unreliable. This suggests that nearly half of the coding in the hospital was found to be untrustworthy, indicating room for improvement in ensuring consistent coding practices.

Table 2. Completeness Variable

Variable Completeness	Freq	Percentage
Complete	42	70
Not Complete	18	30
Total	60	100

Based on the data provided in Table 2, 70% (42 cases) of the coding data were complete, while 30% (18 cases) were incomplete. This indicates a relatively high level of data completeness, which is crucial for maintaining accurate medical records. However, the 30% incompleteness rate shows a significant gap that could affect the overall quality of patient information and care management.

Table 3. Timeliness Variable

Variable Timeliness	Freq	Percentage
Timelines	39	65
Not Timeliness	21	35
Total	60	100

Based on the information provided in Table 3, 65% (39 cases) of the coding was completed promptly, while 35% (21 cases) was not. Timeliness is critical for ensuring that diagnoses and treatment plans are swiftly reflected in patient records. The fact that over a third of the records were delayed may hinder the efficiency of hospital operations and patient care.

Table 4. Accuracy Variable

Variable Accuracy	Freq	Percentage
Accuracy	26	43.3
Not Accurate	34	56.7
Total	60	100

Based on the information provided in Table 4, 43.3% (26 cases) of the coding was found to be accurate, while 56.7% (34 cases) were inaccurate. This is a concerning result, as most coding practices were incorrect, which could directly affect the quality of clinical decisions, reimbursement processes, and overall patient outcomes.

Table 5. Relevancy Variable

Variable Relevancy	Freq	Percentage
Relevant	52	86.7
Not Relevant	8	13.3
Total	60	100

Based on the information in Table 5, 86.7% (52 cases) of the coding was relevant, with only 13.3% (8 cases) being classified as irrelevant. This demonstrates a strong correlation between the coding and the cases being handled, indicating that the codes applied were generally applicable and aligned with the clinical context.

Table 6. Legibility Variable

Variabel Legibility	Freq	Percentage
Legibility	47	78.3
Not Legibility	13	21.7
Total	60	100

Based on the information in Table 6, 78.3% (47 cases) of the coding were legible, while 21.7% (13 cases) were not. While most records were readable, a notable portion of data could lead to misinterpretation or errors in patient care due to illegible records.

The following are the factors that affect the quality of code from the seven attributes:

Table 8. Results from the SPSS Test

Variable	Category	Freq	Reliable	Complete	Timeliness	Accuracy Sig 2 sided	Relevancy	Definition	Legibility
Reliable	Reliable								
	Not reliable		-	0.259	0.778	**0.001**	0.454	0.349	0.547
Complete	Complete								
	Not Complete		0.259	-	1.00	0.261	1.00	**0.001**	**0.046**
Timelines	Timelines				-				
	Not Timelines		0.778	1.00		0.785	1.00	0.737	1.00
Accuracy	Accurate								
	Not Accurate		**0.001**	**0.261**	0.785	-	1.00	1.00	0.529
Relevancy	Relevant								
	Not Relevant		0..454	1.00	1.00	1.00	-	1.00	0.182
Legibility	Legibility								
	Not Legibility		0.547	**0.046**	1.00	0.529	0.182	1.00	-

The results in Table 8 show that a good data source can support other coding variables, such as coding accuracy, reliability or consistency, completeness of data sources, understanding of the writing in the medical record, and legibility.

Table 9. Odds Ratio Variable

Variable	OR	CI 95 for OR
Accuracy*Reliable	8.782	2.651-29.491
Legibility*Complete	3.818	1.059-13.768

The results of this study are variables that affect the quality of coding, such as coding accuracy, reliability or consistency, completeness of data sources, understanding of the writing in the medical record, and legibility. The odds ratio values indicate that consistency can enhance the accuracy of clinical codes by a factor of 8. Conversely, a lack of consistency can decrease accuracy by the same factor. The odds ratio (OR) analysis revealed that higher consistency significantly increases the likelihood of accurate clinical coding, with consistent coding practices enhancing coding accuracy by a factor of eight compared to inconsistent practices. Specifically, the odds of achieving accurate coding were eight times greater when coding was performed consistently than when it was not. Conversely, a lack of consistency was associated with an eightfold reduction in the odds of accurate coding. These findings underscore the critical role of ensuring standardized, reliable documentation and coding processes to maintain high-quality clinical data and reduce the risk of systematic errors in health information systems.

DISCUSSION

While some researchers have primarily focused on assessing clinical coding accuracy, a more comprehensive approach emphasizes that the standards for evaluating coding quality should encompass both accuracy and completeness. The literature distinguishes completeness and accuracy as two separate but equally critical dimensions of coding quality [6,11,19]. Accuracy pertains to the correct assignment of codes that precisely reflect the documented diagnoses and procedures, while completeness refers to the extent to which all relevant diagnoses, procedures, and clinical details are captured in the coding process. Studies by Jordan et al. and others [21] [23,24] have explicitly defined these dimensions, setting clear standards for what constitutes correct and complete morbidity coding. Their findings demonstrated that neglecting either dimension could significantly distort clinical data quality and health service evaluation [25]. Moreover, several other studies have reinforced the idea that both completeness and accuracy are pivotal in determining overall coding quality. For instance, recent evidence suggests that automated coding systems, when properly implemented, tend to produce coding that is both more complete and more accurate than manual coding processes. Automation helps minimize human errors such as oversight, misinterpretation, and fatigue, common in manual coding environments. However, despite the growing body of evidence, some research evaluates clinical coding quality without fully integrating these dual quality-measuring elements, thus potentially overlooking the nuanced interplay between correctness and completeness in influencing health data validity. Integrating both aspects in coding audits is essential for producing reliable epidemiological data, ensuring proper healthcare funding, and supporting accurate public health surveillance.

Other studies have consistently demonstrated that completeness and legibility are essential components of clinical documentation quality, significantly influencing the accuracy of clinical coding [26,27]. Completeness ensures that all relevant clinical information is captured, while legibility enables coders to interpret and code the information provided correctly. In this study, consistency was also found to have a statistically significant impact on coding accuracy ($p = 0.001$), indicating that consistent documentation practices enhance the likelihood of accurate clinical coding. Similarly, the completeness of the data sources significantly influenced the understanding of medical record content, including the interpretation of abbreviations ($p = 0.001$), further reinforcing the interconnectedness of documentation quality elements. Completeness was additionally found to affect the readability of information in medical records, and vice versa, with this relationship being statistically significant (p = 0.046). These findings are aligned with previous research suggesting that the frequent use of abbreviations can increase the risk of misinterpretation, leading to procedural coding errors and decreased coding reliability [12,20]. Regarding survival analysis, a study conducted in the United States showed a ninety percent coding completeness rate [6,12,28,29]. Furthermore, survival analysis studies have shown that coding completeness is a critical factor in clinical outcomes research; for instance, a study conducted in the United States reported a coding completeness rate of 90%, highlighting the achievable benchmarks in documentation and coding practices when rigorous standards and systematic approaches are applied. Together, these findings underscore the necessity of improving documentation completeness, legibility, and consistency as foundational strategies to enhance clinical coding quality and overall healthcare data integrity.

The understanding of written clinical material, encompassing both complete terminologies and their abbreviated forms, has been shown to significantly impact the ease of information retrieval and interpretation. Specifically, enhanced comprehension of written documentation was associated with a threefold increase in the ease of reading and understanding medical records. Clear and standardized language use minimizes ambiguity, reduces coder misinterpretation, and facilitates more accurate clinical coding practices. The American Health Information Management Association (AHIMA) has emphasized that collecting accurate and complete health data is fundamental to healthcare delivery and clinical research, public reporting, reimbursement systems, and evidence-based policymaking. To uphold the integrity of coded data and ensure its effective transformation into actionable information, AHIMA stresses that all users—including clinicians, coders, and data managers—must consistently apply standardized coding rules, conventions, guidelines, and definitions[11,29–33]. Consistency in documentation and coding practices is therefore critical in creating reliable healthcare datasets, promoting interoperability across health information systems, and supporting informed clinical and administrative decision-making. In the absence of uniform adherence to these standards, variations in data quality could undermine patient safety initiatives, resource allocation strategies, and the broader objectives of health system improvement.

CONCLUSIONS AND RECOMMENDATIONS

In conclusion, among the six variables examined, reliability, completeness, timeliness, accuracy, relevancy, and legibility, the most significant factors influencing coding quality were definition and completeness, followed by accuracy and reliability. The highest coding quality levels were observed in definition and completeness, highlighting their central role in ensuring accurate clinical documentation and coding. Based on these findings, future research is recommended to further test these six variables by focusing on specific disease categories, employing a larger sample size, and implementing more structured audit procedures. Additionally, incorporating validation through expert review by clinical coding specialists is suggested to minimize potential bias and enhance the validity of audit results.

ACKNOWLEDGEMENT

Thank you to all respondents involved in this research and to APTIRMIKI for funding it.

REFERENCES

[1] Eckert AJ, Fritsche A, Icks A, Siegel E, Mueller-Stierlin AS, Karges W, et al. Common procedures and conditions leading to inpatient hospital admissions in adults with and without diabetes from 2015 to 2019 in Germany. Wiener Klinische Wochenschrift 2023;135:325–35. https://doi.org/10.1007/s00508-023-02153-z.

[2] Pediatrics AA of. Endocrinology: Updated ICD-10-CM Diabetes and Hypoglycemia Codes. AAP Pediatric Coding Newsletter 2024;20:3–4. https://doi.org/10.1542/pcco_book246_document001.

[3] Daniels A, Muloiwa R, Myer L, Buys H. Examining the reliability of ICD-10 discharge coding in Red Cross War Memorial Children's Hospital administrative database. South African Medical Journal 2021;111:137–42. https://doi.org/10.7196/SAMJ.2021.V111I2.14686.

[4] Kaur R, Ginige JA, Obst O. A Systematic Literature Review of Automated ICD Coding and Classification Systems using Discharge Summaries 2021.

[5] Halid M. Hubungan Kelengkapan Informasi Klinis dengan Ketepatan Kode Diagnosis Berdasarkan ICD-10 pada Kasus Fracture. Jurnal Manajemen Informasi Kesehatan Indonesia 2024;12. https://doi.org/10.33560/jmiki.v12i2.542.

[6] Birmpili P, Atkins E, Li Q, Johal AS, Waton S, Williams R, et al. Evaluation of the ICD-10 system in coding revascularisation procedures in patients with peripheral arterial disease in England: A retrospective cohort study using national administrative and clinical databases. EClinicalMedicine 2023;55.

[7] Nelson SJ, Yin Y, Trujillo Rivera EA, Shao Y, Ma P, Tuttle MS, et al. Are ICD codes reliable for observational studies? Assessing coding consistency for data quality. Digit Health 2024;10:20552076241297056. https://doi.org/10.1177/20552076241297056.

[8] Roba RB, Letta MU, Aychiluhim TN, Minneeneh GA. Intake, digestibility, growth performance and blood profile of rams fed sugarcane bagasse or rice husk treated with *Trichoderma viride* and effective microorganisms. Heliyon 2022;8:e11958. https://doi.org/10.1016/j.heliyon.2022.e11958.

[9] Li M, Fei Z, Zeng M, Wu F-X, Li Y, Pan Y, et al. Automated ICD-9 coding via a deep learning approach. IEEE/ACM Transactions on Computational Biology and Bioinformatics 2018;16:1193–202.

[10] Atolagbe OO, Romano PS, Southern DA, Wongtanasarasin W, Ghali WA. Coding rules for uncertain and "ruled out" diagnoses in ICD-10 and ICD-11. BMC Medical Informatics and Decision Making 2024;21:386. https://doi.org/10.1186/s12911-024-02661-6.

[11] Horsky J, Drucker EA, Ramelson HZ. Accuracy and Completeness of Clinical Coding Using ICD-10 for Ambulatory Visits. AMIA . Annual Symposium Proceedings AMIA Symposium 2018;2018:912–20.

[12] Shepheard J. Clinical coding and the quality and integrity of health data. HIM J 2020;49:3–4. https://doi.org/10.1177/1833358319874008.

[13] Gao S, Fang A, Huang Y, Giunchiglia V, Noori A, Schwarz JR, et al. Empowering biomedical discovery with AI agents. Cell 2024;187:6125–51. https://doi.org/10.1016/j.cell.2024.09.022.

[14] Garza MY, Williams T, Ounpraseuth S, Hu Z, Lee J, Snowden J, et al. Error rates of data processing methods in clinical research: A systematic review and meta-analysis of manuscripts identified through PubMed. International Journal of Medical Informatics 2025;195:105749. https://doi.org/10.1016/j.ijmedinf.2024.105749.

[15] Catling F, Spithourakis GP, Riedel S. Towards automated clinical coding. International Journal of Medical Informatics 2018;120:50–61. https://doi.org/10.1016/j.ijmedinf.2018.09.021.

[16] Dangal A, Maharjan S. Clinical audit of computed tomography requisition form in a Himalayan country, Nepal. Journal of Medical Imaging and Radiation Sciences 2022;53:S11–2. https://doi.org/10.1016/j.jmir.2022.10.038.
[17] Khurram SA, Warner C, Henry AM, Kumar A, Mohammed-Ali RI. Accuracy of clinical coding for procedures in oral and maxillofacial surgery. British Journal of Oral and Maxillofacial Surgery 2016;54:894–7. https://doi.org/10.1016/j.bjoms.2016.05.031.
[18] Dong H, Suárez-Paniagua V, Whiteley W, Wu H. Explainable automated coding of clinical notes using hierarchical label-wise attention networks and label embedding initialisation. Journal of Biomedical Informatics 2021;116:103728. https://doi.org/10.1016/j.jbi.2021.103728.
[19] Hamid Moghaddasi, Reza Rabie, Nasrin Sadeghi. Improving the quality of clinical coding: a comprehensive audit model 2014. https://jhmi.sums.ac.ir/article_42618.html (accessed September 15, 2022).
[20] Abdulla S, Simon N, Woodhams K, Hayman C, Oumar M, Howroyd LR, et al. Improving the quality of clinical coding and payments through student doctor–coder collaboration in a tertiary haematology department. BMJ Open Quality 2020;9:e000723.
[21] Stausberg J, Lehmann N, Kaczmarek D, Stein M. Reliability of diagnoses coding with ICD-10. International Journal of Medical Informatics 2008;77:50–7.
[22] Alonso V, Santos JV, Pinto M, Ferreira J, Lema I, Lopes F, et al. Problems and barriers during the process of clinical coding: a focus group study of coders' perceptions. Journal of Medical Systems 2020;44:1–8.
[23] Young M, Smith MA. Standards and Evaluation of Healthcare Quality, Safety, and Person-Centered Care. StatPearls, Treasure Island (FL): StatPearls Publishing; 2025.
[24] Davies A, Ahmed H, Thomas-Wood T, Wood F. Primary healthcare professionals' approach to clinical coding: a qualitative interview study in Wales. Br J Gen Pract 2025;75:e43–9. https://doi.org/10.3399/BJGP.2024.0036.
[25] Cotton I, McWherter B, Tenbrink T, Sherren K. Comparing thematic and search term-based coding in understanding sense of place in survey research. Journal of Environmental Psychology 2024;96:102339. https://doi.org/10.1016/j.jenvp.2024.102339.
[26] Gupta A, Brazier B, Rengarajan L, Narendran P, Kempegowda P. Determining factors influencing hospital stay for individuals admitted with diabetes-related ketoacidosis – findings from DEKODE length of stay quality improvement project. Clinical Medicine 2024;24:100255. https://doi.org/10.1016/j.clinme.2024.100255.
[27] Atkin C, Khosla R, Belsham J, Hegarty H, Hennessy C, Sapey E. Strategies to identify medical patients suitable for management through same-day emergency care services: A systematic review. Clinical Medicine 2024;24:100230. https://doi.org/10.1016/j.clinme.2024.100230.
[28] Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol 2015;7:449 90. https://doi.org/10.2147/CLEP.S91125.
[29] Wilson R, O'Neil ME, Ntekop E, Zhang K, Ren Y. Coding Completeness and Quality of Relative Survival-Related Variables in the National Program of Cancer Registries Cancer Surveillance System 1995–2008. J Registry Manag 2014;41:65 97.
[30] Adeleke IT, Ajayi OO, Jimoh AB, Adebisi AA, Omokanye SA, Jegede MK. Current Clinical Coding Practices and Implementation of ICD-10 in Africa: A Survey of Nigerian Hospitals. American Journal of Health Research 2014;3:38. https://doi.org/10.11648/j.ajhr.s.2015030101.16.
[31] White RH, Garcia M, Sadeghi B, Tancredi DJ, Zrelak P, Cuny J, et al. Evaluation of the predictive value of ICD-9-CM coded administrative data for venous thromboembolism in the United States. Thrombosis Research 2010;126:61 7.
[32] Perotte A, Pivovarov R, Natarajan K, Weiskopf N, Wood F, Elhadad N. Diagnosis code assignment: models and evaluation metrics. Journal of the American Medical Informatics Association 2014;21:231 7.
[33] Biruk E, Habtamu T, Taye G, Ayele W, Tassew B, Nega A, et al. Improving the Quality of Clinical Coding through Mapping of National Classification of Diseases (NCoD) and International Classification of Diseases (ICD-10). Ethiopian Journal of Health Development 2021;35.