International Seminar and Workshop on Public Health Action
https://swopha.dinus.ac.id

Correlation Between Medical Terminology Accuracy and ICD-10 Coding
A Retrospective Quantitative Study

Gama Bagus Kuntoadi[1*], Sucipto[2], Miftah Parid Firmansyah[3]

1STIKes Widya Dharma Husada Tangerang
2STIKes Widya Dharma Husada Tangerang
3STIKes Widya Dharma Husada Tangerang

Article Info
Article History
Submitted: 13-07-2025
Revised: 05-09-2025
Accepted: 24-08-2025
Keywords:
Medical terminology; ICD-10 code; Corellation; Accuracy, Consistency

Abstract

Background: Accurate medical terminology is essential for diagnostic documentation and forms the basis of reliable ICD-10 coding. Inconsistencies in terminology may lead to coding errors, reduced data quality, and inaccurate reimbursement for health insurance claims.
Methods: This retrospective quantitative study employed a cross-sectional design at Setia Mitra Hospital. Data were analyzed to evaluate the accuracy of medical terminology in diagnosis writing, the accuracy of ICD-10 coding, and the correlation between the two variables.
Results: Out of 137 outpatient diagnoses, 97 (71%) were consistent with standard medical terminology, while 40 (29%) were not. Regarding ICD-10 coding, 65 codes (47.44%) were accurate, whereas 35 (25.54%) were inaccurate. Chi-Square correlation analysis yielded a p-value of 0.276 ($p > 0.05$), indicating no significant association between the accuracy of medical terminology and ICD-10 coding accuracy.
Conclusion: The study found no correlation between medical terminology accuracy and ICD-10 coding accuracy. These findings highlight the need for standardized documentation, continuous training, stronger coder–clinician collaboration, and routine audits to improve diagnostic and coding reliability. Future research may involve multiple facilities, inpatient data, or qualitative approaches to explore contextual factors affecting coding accuracy.

Correspondence Address:
Jl. Pajajaran No.1, Pamulang, Tangerang Selatan, Banten, Indonesia
E-mail: okudagama@gmail.com

eISSN 3063-2439

Introduction

Medical terminology is a specialized field that studies medical terminology used as a means of communication among healthcare professionals (Bagus Kuntoadi et al., 2021). Medical terminology is a complex system that plays an important role in communication, education, and healthcare research (Akbarkhadjaeva, 2020). Accurate and standardized medical terminology is a cornerstone of clinical documentation and a prerequisite for reliable International Classification of Diseases, 10th Revision (ICD-10) coding. High-quality coding is essential not only for ensuring accurate epidemiological data but also for supporting hospital management, quality monitoring, and reimbursement systems, particularly in insurance-based health care such as Indonesia's National Health Insurance (BPJS) program. Conversely, inaccurate or inconsistent terminology may lead to miscoding, reduced data validity, and financial inefficiencies, ultimately affecting both patient care and institutional performance (WHO, 2019) (O'Malley et al., 2005).

Reliable information which can only be derived from accurate data is crucial to the success of the health system. Since encoded data on diagnoses and procedures are put to a broad range of uses, the accuracy of coding is imperative. Accuracy of coding with the International Classification of Diseases, 10th revision (ICD-10) is impeded by a manual coding process that is dependent on the medical records officers' level of experience/knowledge of medical terminologies (Olagundoye et al., 2021).

Completeness of patient medical information in medical records is one indicator in assessing the quality of health services. Complete and accurate medical records contribute to the accuracy of medical staff in making a diagnosis so that they can determine the correct diagnosis code according to the ICD-10 guidelines (Muchlis & Ramadhanty, 2021).

Previous studies have demonstrated persistent challenges in maintaining coding accuracy. In Indonesia, several studies highlight the same issue. A study on upper respiratory infections and hypertension found a significant association between documentation completeness and ICD-10 accuracy ($p = 0.000$) (Muchlis & Ramadhanty, 2021). Another investigation, Sepsis is underreported by ICD-10 coding, which may affect the perceived incidence of sepsis in clinical, claims, and administrative databases. Future research is needed to explore if greater consistency in ICD-10 code definitions and enhanced quality measures for ICD-10 coders can improve the diagnostic accuracy of sepsis from these large databases (Liu et al., 2022).

Despite this evidence, limited research has examined the direct correlation between medical terminology accuracy in diagnostic writing and ICD-10 coding accuracy in Indonesian hospital settings. To address this gap, the present study applies a retrospective cross-sectional design to analyze outpatient documentation at Setia Mitra Hospital. The study specifically aims to (i) assess the accuracy of medical terminology, (ii) evaluate ICD-10 coding accuracy, and (iii) examine the correlation between both variables. These research hypothesis are:

Null Hypothesis (H_0): there is no significant correlation between the accuracy of medical terminology in diagnosis documentation and the accuracy of ICD-10 coding. Alternative Hypothesis (H_1): there is a significant correlation between the accuracy of medical terminology in diagnosis documentation and the accuracy of ICD-10 coding.

Methods

Study Design and Setting

This study employed a retrospective observational design with a cross-sectional approach. The research was conducted in the Medical Records Unit of Setia Mitra Hospital, located in South Jakarta, Indonesia.

Population and Sample

The study population comprised all outpatient medical records at Setia Mitra Hospital in 2023, totaling 82,937 records. The required sample size was calculated using Slovin's formula at a 90% confidence level, resulting in 100 medical records. From 100 medical records, 137 diagnoses were identified, including both primary and secondary diagnoses. A systematic random sampling technique was applied with an interval of 829 to ensure representativeness of the outpatient population.

Variables

The primary variables were Independent variable: accuracy of medical terminology in physicians' diagnostic documentation, Dependent variable: accuracy of ICD-10 clinical coding assigned to the respective diagnoses.

Data Collection

Data were collected through structured observation using an observation checklist developed according to ICD-10 guidelines and hospital documentation standards. The checklist evaluated the correctness of medical terminology and the appropriateness of corresponding ICD-10 codes. Prior to analysis, all data underwent completeness and validity checks, followed by coding, data entry, and cleaning.

Data Analysis

Data analysis was conducted using the latest version of IBM SPSS Statistic software (IBM Corp). Descriptive statistics were employed to present the distribution of medical terminology and ICD-10 coding accuracy. For inferential analysis, a Chi-Square test was performed to examine the correlation between the accuracy of medical terminology and ICD-10 codes. Statistical significance was set at $p < 0.05$.

Ethical Considerations

Patient confidentiality was maintained by anonymizing all medical records prior to data analysis.

Results

Univariate Analysis

Table 1 presents the distribution of medical terminology accuracy based on 137 outpatient diagnoses (including primary and secondary diagnoses) at Setia Mitra Hospital.

Table 1. Distribution of Medical Terminology Accuracy in Outpatient Diagnoses.

No	Accuracy of Medical Terminology	Frequency (f)	Percentage (%)
1	Accurate	97	71.0
2	Inaccurate	40	29.0
Total		137	100.0

From Table 1, 71% of outpatient diagnoses were documented using accurate medical terminology, whereas 29% were not. This indicates that while most physicians adhered to standard terminology, nearly one-third of diagnoses contained deviations from proper medical terminology, potentially affecting subsequent ICD-10 coding.

Table 2 shows the accuracy of ICD-10 coding assigned to the same outpatient diagnoses.

ICD-10 Coding Accuracy	Frequency (f)	Percentage (%)
Accurate	65	47.4
Inaccurate	35	25.5
Not Coded	37	27.0
Total	137	100.0

As shown in Table 2, less than half of the diagnoses (47.4%) were coded accurately, while 25.5% were miscoded and 27.0% were not coded at all. The proportion of uncoded diagnoses is noteworthy, suggesting potential gaps in documentation practices or coder workload. Overall, these findings highlight variability in both medical terminology usage and ICD-10 coding quality in the outpatient setting.

Bivariate Analysis

Bivariate analysis was conducted to examine the correlation between the accuracy of medical terminology and the accuracy of ICD-10 coding in outpatient diagnoses at Setia Mitra Hospital. The results of multivariate analysis using Pearson Chi-Square analysis with the SPPS application to analyze the correlation between the appropriateness of medical terminology use and the accuracy of ICD-10 codes at Setia Mitra Hospital. The Chi-Square test was applied with a significance level of $p < 0.05$.

Table 3. Terminologi Medis * kode ICD 10 Crosstabulation

			kode ICD 10		Total
			Tidak Tepat	Tepat	
Terminologi Medis	Tidak Akurat	Count	11	14	25
		Expected Count	8,8	16,3	25,0
		% within Terminologi Medis	44,0%	56,0%	100,0%
		% within kode ICD 10	31,4%	21,5%	25,0%
		% of Total	11,0%	14,0%	25,0%
	Akurat	Count	24	51	75
		Expected Count	26,3	48,8	75,0
		% within Terminologi Medis	32,0%	68,0%	100,0%
		% within kode ICD 10	68,6%	78,5%	75,0%
		% of Total	24,0%	51,0%	75,0%
Total		Count	35	65	100
		Expected Count	35,0	65,0	100,0
		% within Terminologi Medis	35,0%	65,0%	100,0%
		% within kode ICD 10	100,0%	100,0%	100,0%
		% of Total	35,0%	65,0%	100,0%

Table 4. Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,187[a]	1	,276		
Continuity Correction[b]	,718	1	,397		
Likelihood Ratio	1,162	1	,281		
Fisher's Exact Test				,335	,198
Linear-by-Linear Association	1,175	1	,278		
N of Valid Cases	100				
a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 8,75.					
b. Computed only for a 2x2 table					

Table 5. Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Contingency Coefficient	,108	,276
N of Valid Cases		100	

Chi-Square test, p = 0.276 (p > 0.05). According to the Table 4, the results of the assumption testing using the Pearson Chi-square test yielded a p-value of 0.276, meaning p > 0.05, so H0 is accepted, meaning that there is no correlation between the accuracy of medical terminology used in diagnoses and the accuracy of ICD-10 codes for patient diagnoses at Setia Mitra Hospital.

These findings suggest that although accurate terminology may facilitate ICD-10 coding, it does not guarantee coding accuracy. Other factors—such as coder expertise, documentation completeness, and institutional coding practices—likely play an important role in determining coding accuracy.

Discussion

This study evaluated the accuracy of medical terminology in outpatient diagnoses, ICD-10 coding accuracy, and the correlation between both variables at Setia Mitra Hospital. The findings revealed that while the majority of diagnoses (71%) were written using accurate medical terminology, less than half (47.4%) of the ICD-10 codes were accurate, and 27% of diagnoses were not coded at all. Furthermore, no statistically significant correlation was identified between medical terminology accuracy and ICD-10 coding accuracy (p = 0.276).

These results are contradict with a study in Bengkulu Province Indonesia, which reported the accuracy of medical terminology can affect the accuracy of inpatient diagnosis codes, as evidenced by 49 (52.7%) medical records with accurate medical terminology resulting in accurate codes, however, 6 (6.5%) medical records with accurate medical terminology but inaccurate diagnosis codes, and 38 (40.8%) medical records with inaccurate medical terminology resulting in inaccurate codes. The statistical test used to test the hypothesis was the Chi-square Test with Ha, which found a relationship between the accuracy of medical terminology and the accuracy of coding inpatient cases at RSKJ Soeprapto in Bengkulu Province in 2020. The p-value in this study was 0.000 with a significance level of 0.05. This indicates that H0 is rejected or Ha is accepted, thus concluding that there is a relationship between the accuracy of medical terminology writing and the accuracy of coding in inpatient cases at RSKJ Soeprapto Hospital in Bengkulu Province with a p-value ≤ α 0.005. The accuracy of medical terminology can influence the accuracy of inpatient case coding by healthcare worker (Heltiani et al., 2022). The relatively high accuracy of medical terminology suggests that most physicians adhere to standardized documentation practices. However, the considerable proportion of inaccurate or missing ICD-10 codes highlights systemic challenges in coding implementation.

In Indonesia, several studies have emphasized the critical role of complete and standardized documentation in improving coding accuracy. Based on previous research entitled “Evaluation of the Accuracy of Medical Terminology and Its Relationship to the Accuracy of Clinical Coding in Health Facilities: Systematic Literature Review,” it appears that the average accuracy rate of clinical coding in health facilities is 48%, with 14 articles having accuracy rates above or below 50%. The average accuracy rate of medical terminology in healthcare facilities is 55%, with 11 articles having an accuracy rate below 50%. Up to 79% of the findings in this study indicate a relationship between the accuracy of clinical coding and the accuracy of medical terminology (Firdayana et al., 2023). Meanwhile, in a study titled “Analysis of the Accuracy of Medical Terminology for Respiratory System Diseases According to the ICD-10 Disease

Classification at the Grogol Petamburan Subdistrict Health Center," the accuracy of medical terminology usage in medical records was identified as 74 (82%) and the inaccuracy of medical terminology usage as 16 (18%) due to the use of abbreviations, Indonesian language, and spelling/typing errors (Irawan et al., 2022).

The present study did not find such an association, suggesting that factors beyond terminology accuracy—such as coder expertise, documentation completeness, institutional coding policies, and workload distribution—may exert stronger influence on coding quality.

The absence of correlation in this study underscores the complexity of coding accuracy as a multifactorial outcome. While accurate terminology is necessary, it may not be sufficient without parallel investments in coder training, interprofessional collaboration, and routine auditing systems. Hospitals may therefore benefit from implementing structured coding audits, feedback mechanisms, and targeted training programs to bridge the gap between clinical documentation and coding practices.

Strengths and Limitations: The use of real-world hospital data and a standardized checklist are strengths of this study. However, limitations include its single-center design, restriction to outpatient cases, and the retrospective nature of the analysis. These factors may limit the generalizability of findings.

Implications and Future Research: The results suggest that improving ICD-10 coding accuracy requires not only accurate terminology but also organizational support and coder–clinician collaboration. Future research should extend to multiple healthcare facilities, incorporate inpatient data, and apply mixed-method approaches, including qualitative interviews with clinicians and coders, to better capture contextual factors affecting coding practices.

Conclusion

This study assessed the relationship between the accuracy of medical terminology and ICD-10 coding at Setia Mitra Hospital. Although 71% of diagnoses were documented using accurate medical terminology, only 47.4% of ICD-10 codes were correct, and 27% of diagnoses were not coded. Statistical analysis revealed no significant correlation between medical terminology accuracy and ICD-10 coding accuracy ($p > 0.05$). These findings indicate that while standardized terminology is essential, coding accuracy depends on additional factors such as coder expertise, documentation completeness, and institutional coding practices. Hospitals should therefore prioritize routine audit systems, enhance coder–clinician collaboration, and implement continuous training to improve coding quality and ensure reliable health information systems. Future research should expand to multicenter settings, explore both outpatient and inpatient data, and incorporate qualitative methods to capture contextual barriers and enablers of coding accuracy.

Author Contributions

GBK, MPF collated and analyzed data to prepare the manuscript. S, and MPF contributed data and reviewed the manuscript for revisions. GBK performed the final review and endorse the paper for publication.

Funding

The author(s) received financial support for the research, authorship,and/or publication of this article from internal research grant fund from STIKes Widya Dharma Husada Tangerang.

Institutional Review Board Statement

Ethical approval was not sought for the present study because the institution either the Setia Mitra hospital and STIKes Widya Dharma Husada Tangerang does not require ethical approval for reporting this research.

Acknowledgments

The authors would like to thank the committee of the 5th International Seminar and Workshop on Public Health Action (ISWOPHA 2025) for the opportunity and support in this publication process. The authors is also grateful for the support from Setia Mitra Hospital, and STIKes Widya Dharma Husada Tangerang, Indonesia, as an educational institution which have provided moral and material support in implementing the Tridharma of Higher Education.

References

Akbarkhadjaeva, F. (2020). THE NATURE OF MEDICAL TERMINOLOGY. Theoretical & Applied Science, 86(06), 432–434. https://doi.org/10.15863/TAS.2020.06.86.80

Bagus Kuntoadi, G., Kristina, I., & Agustini, H. (2021). Buku Ajar Terminologi Medis untuk Mahasiswa Rekam Medis & Infokes (G. Bagus Kuntoadi (ed.)). Insan Cendekia Mandiri.

Firdayana, F., Nurmalasari, M., Hosizah, H., & Afrizal, S. H. (2023). Evaluation of the Accuracy of Medical Terminology and Its Relationship to the Accuracy of Clinical Coding in Health Facilities: Systematic Literature Review. Jurnal Ners, 7(2), 1171–1181. https://doi.org/10.31004/jn.v7i2.16381

Heltiani, N., Manalu, A. K., & Anggita, F. (2022). Analisis Ketepatan Penulisan Terminologi Medis terhadap Keakuratan Kodefikasi Kasus Rawat Inap RSKJ Soeprapto Provinsi Bengkulu. SEMINAR NASIONAL PENELITIAN DAN PENGABDIAN 2022.

Irawan, F., Deasy Rosmala Dewi, Nanda Rumana, & Muniroh. (2022). Analisis Ketepatan Istilah Terminologi Medis Penyakit Sistem Respirasi Sesuai Klasifikasi Penyakit ICD-10 di Puskesmas Kecamatan Grogol Petamburan. SEHATMAS: Jurnal Ilmiah Kesehatan Masyarakat, 1(2), 230–239. https://doi.org/10.55123/sehatmas.v1i2.227

Liu, B., Hadzi-Tosev, M., Liu, Y., Lucier, K. J., Garg, A., Li, S., Heddle, N. M., Rochwerg, B., & Ning, S. (2022). Accuracy of International Classification of Diseases, 10th Revision Codes for Identifying Sepsis: A Systematic Review and Meta-Analysis. Critical Care Explorations, 4(11), e0788. https://doi.org/10.1097/CCE.0000000000000788

Muchlis, A., & Ramadhanty, R. A. (2021). Correlation Between Medical Information Completeness and Accuracy of The Diagnosis Code for Upper Respiratory Tract Infection and Hypertension Based on ICD-10 in Medical Record at Cibening Health Center. Muhammadiyah Medical Journal, 2(1), 1. https://doi.org/10.24853/mmj.2.1.1-6

O'Malley, K. J., Cook, K. F., Price, M. D., Wildes, K. R., Hurdle, J. F., & Ashton, C. M. (2005). Measuring Diagnoses: ICD Code Accuracy. Health Services Research, 40(5p2), 1620–1639. https://doi.org/10.1111/j.1475-6773.2005.00444.x

Olagundoye, O., van Boven, K., Daramola, O., Njoku, K., & Omosun, A. (2021). Improving the accuracy of ICD-10 coding of morbidity/mortality data through the introduction of an electronic diagnostic terminology tool at the general hospitals in Lagos, Nigeria. BMJ Open Quality, 10(1), e000938. https://doi.org/10.1136/bmjoq-2020-000938

WHO. (2019). International Classification of Diseases (ICD-10). World Health Organization (W.H.O).