Development of Diagnostic Terminology and Coding for General Practice and Empirical Study of Its Application: A Postprint

Authors: Hong Yuchun 1, Wu Hua 2, Du Yishan 3, Li Shuran 4, Sun Wenmin 1, Ye Mingyu 5, Zhang Yongjian 6, Li Yang 1*

Date: 2023-07-26T00:00:00+00:00

Abstract

Background Establishing a unified set of general practice diagnostic terminology and codes is an important means to improve the service level of community health service institutions and promote the construction of a general practice system in China. However, there has been a lack of standardized diagnostic terminology and codes suitable for the current development status of general practice in China, leading to non-standardized and non-uniform diagnoses by general practitioners, which has seriously hindered the development of general practice in China.

Objective To construct a set of general practice diagnostic terminology and codes for community health services, and to validate the effectiveness of the diagnostic terminology and code set, thereby promoting the standardization and normalization of community general practice diagnosis.

Methods Through literature review and expert consultation, combined with the spectrum of common health problems and disease patterns in communities, a set of general practice diagnostic terminology and codes applicable to community health services was constructed based on ICD-10. From September 2021 to February 2022, an empirical study was conducted by piloting the developed general practice diagnostic terminology and code set in 259 community health service centers across two districts in Shenzhen. Using backend diagnostic data from the community health service information system and questionnaire surveys of community general practitioners as the primary data sources, the implementation of the pilot was analyzed to validate the rationality and effectiveness of the diagnostic terminology and code set.

Results The compilation of the general practice diagnostic terminology and code set was completed, comprising 41 categories of general practice diagnoses with

6,629 standardized general practice diagnostic terms and codes. The pilot results from the two Shenzhen districts showed that the proportion of standardized diagnostic names in the standardized intervention area was 86.13%–95.38% higher than in the non-pilot area; the diagnostic names used in the pilot area during this period decreased by 96.60% year-over-year, with the standardized intervention significantly focusing commonly used general practice diagnoses that were largely consistent with the community disease spectrum, demonstrating high aggregation of diagnoses for common community diseases. Questionnaire survey results showed that 87.31% (227/260) of respondents believed the code set promoted standardization of community diagnosis, and 77.31% (201/260) supported its implementation.

Conclusion This study preliminarily established a standardized and normalized set of general practice diagnostic terminology and codes, forming a community general practice diagnostic terminology dictionary that solves the problem of multiple terms for a single meaning in data analysis, while also accommodating referral systems and medical insurance integration, effectively improving the efficiency and standardization level of community diagnosis and treatment, and positively promoting the development of primary healthcare.

Full Text

Development of General Practice Diagnostic Terminology and Coding and Empirical Study on Its Application

HONG Yuchun[1], WU Hua[2], DU Yishan[3], LI Shuran[4], SUN Wenmin[1], YE Mingyu[5], ZHANG Yongjian[6], LI Yang[1]*

[1]Shenzhen Health Capacity Building and Continuing Education Center, Shenzhen 518000, China
[2]Bao'an People's Hospital, Shenzhen 518000, China
[3]International Centre for Higher Education Innovation under the Auspices of UNESCO, Shenzhen 518000, China
[4]Public Health Service Center of Bao'an District, Shenzhen 518000, China
[5]Futian Management Center of Community Health Service, Shenzhen 518000, China
[6]Peking University Shenzhen Hospital, Shenzhen 518000, China

Corresponding author: LI Yang, Assistant researcher; E-mail: 15816883396@139.com

HONG Yuchun and WU Hua are co-first authors

Abstract

Background: Establishing a unified set of diagnostic terms and codes for general practice is crucial for improving service levels at community health institutions and advancing China's general practice system. However, a lack of

standardized diagnostic terminology and codes tailored to China's general practice development has resulted in inconsistent and non-standard diagnoses among general practitioners, severely impeding the field's progress.

Objective: To construct and validate a standardized diagnostic terminology and coding system for community-based general practice, thereby promoting standardization of community diagnoses.

Methods: Through literature review and expert consultation, we developed a general practice diagnostic terminology and coding system based on ICD-10 and aligned with community health problem and disease spectra. From September 2021 to February 2022, we conducted a pilot study across 259 community health service centers in two Shenzhen districts. Using backend diagnostic data from the community health information system and questionnaire surveys of general practitioners, we analyzed implementation outcomes to verify the system's rationality and effectiveness.

Results: The completed system comprised 41 diagnostic categories with 6,629 standardized terms and codes. Pilot results showed standardized diagnostic name composition in intervention areas was 86.13%–95.38% higher than in non-pilot areas. Diagnostic names used in pilot districts decreased by 96.60% year-over-year, with standardized interventions significantly focusing common diagnoses that aligned with community disease spectra. Survey results indicated 87.31% (227/260) of respondents believed the coding system promoted diagnostic standardization, and 77.31% (201/260) supported its implementation.

Conclusion: This study establishes a preliminary standardized diagnostic terminology and coding system for general practice, creating a community diagnostic dictionary that resolves polysemy issues in data analysis while facilitating referrals and medical insurance integration. The system effectively improves community healthcare efficiency and standardization, positively impacting primary care development.

Keywords: General practice; Community health services; Diagnosis; Terminology coding

---

Introduction

As community health "gatekeepers," primary healthcare institutions bear responsibility for first-contact care and represent a critical component of tiered diagnosis and treatment systems. China's deepening healthcare reform and resource decentralization have driven rapid growth in community patient volumes, while two-way referral channels have increased patient transfers between levels. However, the persistent lack of standardized diagnostic terminology and codes suited to China's general practice context has led to non-standard and inconsistent diagnoses among community general practitioners, severely hindering general practice development. Establishing a unified general practice diagnostic

terminology and coding system would not only facilitate information exchange and sharing between community health institutions and higher-level hospitals but also enable disease statistics for general practice, laying foundations for evaluating practitioner competencies, conducting epidemiological surveys, and monitoring population health. This represents an important pathway for improving community health service levels and developing the general practice system.

The International Classification of Diseases (ICD) is the globally recognized standard for health information. ICD-10, the tenth revision, is currently the most widely used disease classification system worldwide, characterized by its detail, consistency, completeness, and practicality. It enables easy retrieval and implements one-disease-one-code principles while strictly adhering to disease classification principles and evaluation standards, making it a relatively comprehensive international classification system. China has implemented various localized versions, including the National Health Commission's 2012 ICD edition (6-digit extension codes), Beijing's ICD-10 clinical version (6-digit extension codes), and Shanghai's ICD-10 (7-digit detailed codes). Nevertheless, significant controversies remain regarding ICD-10's practical application, necessitating further localization research and empirical validation.

Methods

Literature Review From November 2020 to January 2021, we systematically searched CNKI, VIP, Wanfang Data, and PubMed for Chinese and English literature from the past decade using keywords including "general practice," "diagnostic coding," "diagnostic terminology," "disease classification," and "diagnostic standardization." We also compiled policy documents on diagnostic standardization from various municipalities to establish our methodology for developing the general practice diagnostic terminology and coding system based on ICD-10, clarifying development pathways, work plans, and procedures.

Expert Consultation We established a 36-member expert panel with inclusion criteria of: (1) specialization in general practice, clinical medicine, or medical record management and coding; (2) $ $5 years of work experience; (3) bachelor's degree or higher; (4) intermediate professional title or above; and (5) voluntary participation with relevant domain knowledge. From February to March 2021, we extracted electronic diagnostic data from Shenzhen community health institutions for 2019–2020, including diagnostic names and usage frequencies. After data cleaning to identify different expressions for the same disease and exclude typographical errors and non-standard terminology, we compiled community health problem and disease spectra and mapped them to ICD-10 codes to form a draft comprehensive general practice diagnostic terminology set. We then ranked diagnoses by frequency to create a core terminology draft, retaining all ICD-10 codes as specialist diagnostic terminology.

From April to July 2021, we conducted coding verification through expert group

review by ICD-10 chapter. We established one-disease-one-code relationships between general practice and specialist diagnoses, resolving disputed items through expert discussion. This process yielded the final general practice diagnostic terminology and coding system.

Information System Integration In August 2021, we embedded the general practice diagnostic terminology and coding system into Shenzhen's community health service information platform. We restructured the SOAP (subjective, objective, assessment, plan) interface by dividing the diagnosis section into two columns: general practice diagnosis (left) and specialist diagnosis (right) [Figure 1: see original paper], with specific rules: (1) General practice diagnosis is mandatory and must use standardized ICD-10 codes; specialist diagnosis is optional and allows custom entries. (2) If general practice diagnosis is entered first, specialist diagnosis can be selected from corresponding options, custom-entered, or left blank. If specialist diagnosis is entered first using standardized ICD-10 codes, the general practice diagnosis auto-populates; custom specialist diagnoses require manual general practice diagnosis entry. (3) When saving/suspending/ending consultations, the system prompts "Western medicine diagnosis: General practice diagnosis does not meet requirements, please complete before suspending or ending!" if general practice diagnosis is missing or non-standard.

Empirical Analysis Pilot Implementation: Using cluster sampling, we selected 259 community health service centers in two Shenzhen districts for a six-month pilot (September 2021–February 2022). General practitioners in pilot areas uniformly used the terminology and coding system during daily practice and medical record writing, with dynamic updates based on clinical needs. We standardized data reporting to improve quality. Pilot activities included: restructuring the community health information system to embed the terminology; printing and distributing the *Shenzhen General Practice Diagnostic Classification and Coding Manual (Trial)* with training; establishing pilot safeguard mechanisms and feedback/revision protocols for new codes; forming a pilot working group; and creating communication channels for issue resolution.

Monitoring Indicators: (1) **Applicability and Standardization Validation:** We extracted standardized and non-standardized diagnostic term usage frequencies from the community health information system for pilot areas (October 2021–February 2022) and the same period in the previous year (October 2020–February 2021). (2) **Satisfaction Survey:** Post-pilot questionnaires targeted general practitioners, center managers, and medical record quality control staff in pilot areas. The survey assessed: (a) demographics; (b) satisfaction with the coding system including its impact on standardization (5-point scale) and attitudes toward implementation (5-point scale); and (c) system improvement suggestions. We distributed 260 questionnaires with 100% valid response rate.

Statistical Analysis We used R 3.5.1 software for statistical analysis, presenting categorical data as frequencies and percentages.

Results

Composition of General Practice Terminology and Coding The complete system was derived from ICD-10 and comprised three components: (1) **Core General Practice Diagnoses:** 3,476 high-frequency terms corresponding to common community health problems and disease spectra. (2) **Comprehensive General Practice Diagnoses:** 6,629 terms for non-common problems retained for two-way referral convenience. (3) **Specialist Diagnoses:** All 35,869 ICD-10 terms retained as specialist diagnoses. General practice and specialist codes maintained one-disease-one-code correspondence .

Empirical Study Findings Applicability and Standardization Analysis: (1) During the pilot period (October 2021–February 2022), non-pilot districts used up to 80,807 diagnostic names, with standardized names accounting for only 4.62%–13.87%. In contrast, pilot districts achieved 100% standardized diagnostic name composition, exceeding non-pilot areas by 86.13%–95.38% .

(2) Year-over-year comparison in the two pilot districts showed: total diagnoses decreased by 50.65%; diagnostic names used decreased by 96.60%; and the proportion of top 50 diagnoses increased by 15.92%. The top 10 diagnoses (hypertension, COVID-19 screening, diabetes, hyperlipidemia, health examination, dyspepsia, sleep disorders, type 2 diabetes, chronic gastritis, and essential hypertension) accounted for 39.24% of all diagnoses . Standardization of common disease diagnoses improved substantially. For example, "hypertension" diagnoses consolidated from 2,311 variations (e.g., "hypertension," "grade 2 hypertension," "grade 2 hypertension (very high risk)," "hypertension?," "hypertension medication refill") to 37 standardized terms—a 98.40% reduction .

Satisfaction Survey Results: Among 260 respondents, 209 (80.38%) were general practitioners, 24 (9.23%) were center managers, 17 (6.54%) were quality control staff, and 10 (3.85%) were other personnel. Professional titles included intermediate (64.23%), junior (19.23%), senior (13.08%), and none (3.46%). Results showed 87.31% (227/260) believed the coding system promoted diagnostic standardization, with 70.00% (182/260) reporting improved standardization. While 85.00% (221/260) did not oppose implementation, 77.31% (201/260) actively supported it.

Discussion

Impact on Community Diagnostic Standardization The absence of standardized, efficient primary care terminology and codes has allowed physicians to self-compose diagnostic names, resulting in inconsistent disease descriptions and terminology that compromises clinical information integrity and affects disease

classification, clinical evaluation, and statistical reliability. This study found non-pilot districts using up to 80,000 diagnostic variations, creating significant challenges for data analysis and application.

Our standardized system consolidated diagnoses dramatically—for instance, reducing hypertension-related terms by 98.40%—while focusing common diagnoses to align with community disease spectra. The top 10 diagnoses accounted for 39.24% of total volume post-intervention. Satisfaction surveys revealed 87.31% of respondents recognized the system's standardization benefits. While implementation impacts traditional workflows, these findings underscore the importance of promoting standardized coding awareness among healthcare providers.

ICD-10 as a Foundation for Localized General Practice Coding Medical classification codes enable disease data storage, retrieval, and analysis for mortality statistics, epidemiological surveys, insurance payment, and performance management internationally. The two major systems are WONCA's International Classification of Primary Care (ICPC-2) and ICD-10. ICD-10 emphasizes epidemiology and health management, serving as the globally recognized health information standard.

Most Chinese hospitals currently use ICD-10. However, ICD-10's specialty-based, fine-grained classification can be overly specialized for general practice, which emphasizes comprehensiveness and breadth. Direct application may burden general practitioners with excessive diagnostic options and overlooks social, psychological, and non-disease health issues essential to general practice.

ICPC-2, used in some international primary care settings and previously piloted in Shenzhen, employs a two-axis classification system with coarser granularity than ICD-10's multi-axial approach. Its classification rules include reasons for encounter, process, and diagnosis, imposing high data and operational demands. China's general practice scope differs from other countries, making ICPC-2 unsuitable for current community health institution needs—its coarse granularity cannot accommodate China's large general practice population, and localization is overly complex without established supporting information systems.

This study developed general practice terminology and coding based on ICD-10, community health problem spectra, and disease spectra. Our approach prioritized practical needs, aligning with community health service comprehensiveness through non-uniform granularity—primarily moderate coarseness with appropriate detail for health issues, while considering referrals, specialist 对接, and insurance payment. Pilot results demonstrated significantly improved effectiveness and standardization with high practitioner satisfaction, confirming ICD-10's viability as a theoretical and practical foundation for China's localized general practice coding.

Information System Development for Future Implementation This study demonstrates the necessity of robust information systems for terminology application. Future improvements should focus on personalized enhancements: (1) Implementing logical validation to prevent unnecessary errors during data entry; (2) Customizing functions based on institutional and physician diagnostic patterns, such as prioritizing frequently used diagnoses to improve efficiency and record quality.

In conclusion, we have preliminarily constructed a standardized general practice diagnostic terminology and coding system that unifies multi-source data through information standard 对接, aligning community health information with higher-level hospitals and insurance departments while addressing referrals and payment issues. This system facilitates understanding of primary care disease spectra, guides community healthcare service planning, promotes primary care institution development and general practitioner capacity building, and provides scientific evidence for public health planning and administration.

Limitations: First, the short pilot duration requires further validation of generalizability and replicability through expanded implementation. Second, traditional Chinese medicine diagnoses were not standardized; future work will build on this experience to develop TCM diagnostic standardization.

Author Contributions: HONG Yuchun, WU Hua, and LI Yang conceptualized the study. HONG Yuchun, DU Yishan, and LI Shuran collected and analyzed data and drafted the manuscript. WU Hua and ZHANG Yongjian revised the manuscript. LI Shuran, SUN Wenmin, and YE Mingyu provided guidance and coordination. LI Yang supervised quality control and is responsible for the overall work.

Conflict of Interest: The authors report no conflicts of interest.

References:

[1] SHEN Shuguang, ZHANG Bo. Tiered diagnosis and treatment, primary care first contact, and primary healthcare institution construction[J]. Xuehai, 2016, 27(2): 48-57.

[2] CHEN Minsheng, LAI Wei, ZHANG Zhencheng, et al. Investigation on awareness and willingness of two-way referral in Songjiang District, Shanghai[J]. Journal of Shanghai Jiao Tong University (Medical Edition), 2010, 30(3).

[3] WANG Miao, YU Guangjun, LIU Haifeng, et al. Establishment of disease data quality management system based on ICD-10 coding application[J]. Chinese Hospital Management, 2017, 37(5): 42-45.

[4] ZHOU Jingya, BAI Xue, CUI Shengnan, et al. Systematic evaluation of ICD-10 disease classification coding quality in China[J]. Chinese Hospital Management, 2015, 35(12): 32-35.

[5] HAO Xiaoning, MA Chengyu, LIU Zhiye, et al. Study on satisfaction of both doctors and patients with primary healthcare informatization construction[J]. Health Economics Research, 2020, 37(7): 6-9.

[6] ZHANG Meng, CI Puwa, LIU Haimin, et al. Design and construction of disease classification system for primary healthcare in China[J]. Chinese Medical Record, 2017, 18(2): 33-36. DOI: 10.3969/j.issn.1672-2566.2017.02.013.

[7] XU Wen, TAO Tingping. International Classification of Diseases (ICD-10) and its coding practice[J]. Chinese Journal of Health Statistics, 2003, 20(5): 35-36. DOI: 10.3969/j.issn.1002-3674.2003.05.012.

[8] LÜ Chuanlu, WANG Xia, CAO Hongxin. Reflections on localization research of the 10th revision of International Classification of Diseases[J]. Journal of Naval Medicine, 2016, 37(2): 190-191. DOI: 10.3969/j.issn.1009-0754.2016.02.032.

[9] YANG Xingyu, WANG Henan, YU Lihua, et al. Clinical application and practice of "Chinese Clinical Disease Diagnosis Standard Terminology Set"[J]. Chinese Health Economics, 2021, 40(4): 27-30.

[10] WEN Huilan, LU Yaoneng, LUO Yijuan, et al. Influence of disease classification on promoting tiered diagnosis and treatment system construction[J]. Modern Hospital, 2020, 20(11): 1593-1596. DOI: 10.3969/j.issn.1671-332X.2020.11.011.

[11] ZHANG Xinyi, XIE Ming. Comparison and enlightenment of medical classification coding between China and the United States in healthcare field[J]. Chinese Medical Record, 2018, 19(9): 9-12. DOI: 10.3969/j.issn.1672-2566.2018.09.004.

[12] LU Xiaoqin, LIANG Wannian, JIA Qingchun, et al. International Classification of Primary Care (ICPC) and its application in general practice/family medicine[J]. Chinese General Practice, 2003, 6(1): 86-88. DOI: 10.3969/j.issn.1007-9572.2003.01.049.

[13] ZHANG Meng, CI Puwa, LIU Haimin, et al. Design and construction of disease classification system for primary healthcare in China[J]. Chinese Medical Record, 2017, 18(2): 33-36. DOI: 10.3969/j.issn.1672-2566.2017.02.013.

[14] WANG Wenjun, ZHOU Daoping, XU Shuman, et al. Analysis of influencing factors on quality of front page filling in medical records by resident physicians based on logistic model[J]. Chinese Journal of Health Statistics, 2016, 33(4): 261-263. DOI: CNKI:SUN:ZGWT.0.2016-02-023.

Received: October 26, 2022
Revised: March 21, 2023
Accepted: (Not specified in original)
Edited by: ZHANG Yali