
AI translation · View original & related papers at
chinaxiv.org/items/chinaxiv-202304.00736

Towards a Metadata Model for Clinical Case Records in Traditional Chinese Medicine Diagnosis and Treatment Knowledge Bases: A Postprint

Authors: Li Lu, Ma Jie, Sun Hengyu, Wang Jue

Date: 2023-04-01T16:02:44+00:00

Abstract

[Purpose/Significance] In the management of Traditional Chinese Medicine (TCM) case records, issues such as non-uniform metadata and insufficient precision hinder information retrieval, knowledge sharing, and mining within TCM case record knowledge bases. Researching and establishing a standardized metadata model for TCM case records is crucial for supporting deep multi-dimensional retrieval, knowledge sharing, and mining of TCM case records. [Method/Process] Through literature research methodology, existing TCM-related standards, specifications, and metadata standards were summarized; through web survey methodology, TCM case record instances were collected and, combined with content analysis methodology, the case record content was analyzed to preliminarily define case record metadata elements; subsequently, through field survey methodology, the TCM clinical process was observed and accessed at clinical settings to optimize the preliminarily defined case record metadata elements, yielding the final set of case record metadata elements. [Results/Conclusion] A case record metadata model oriented toward TCM clinical knowledge bases was constructed, which comprehensively describes the TCM clinical process and provides a reference for standardizing TCM case record metadata.

Full Text

Preamble

Volume 65, Issue 2, January 2021
ChinaXiv Collaborative Journal

Research on the Construction of a Medical Case Metadata Model for Traditional Chinese Medicine Diagnosis and Treatment Knowledge Bases

Authors: Li Lu¹, Ma Jie^{1, 2}, Sun Hengyu³, Wang Jue¹

¹ School of Management, Jilin University, Changchun 130022

² Information Resources Research Center, Jilin University, Changchun 130022

³ The First Clinical Hospital of Jilin Academy of Traditional Chinese Medicine, Changchun 130022

Abstract: [Purpose/Significance] In the management of Traditional Chinese Medicine (TCM) medical cases, problems such as inconsistent metadata and insufficient precision hinder information retrieval, knowledge sharing, and mining in TCM case knowledge bases. Establishing a standardized TCM medical case metadata model is crucial for supporting deep, multi-dimensional retrieval, knowledge sharing, and mining of TCM cases. [Method/Process] Through literature review, this study summarized existing TCM-related standards and metadata specifications; through web surveys, it collected TCM case examples and employed content analysis to analyze case contents, preliminarily defining case metadata elements; through field investigations, it observed and interviewed at TCM diagnosis and treatment sites to optimize the preliminarily defined elements, obtaining the final set of case metadata elements. [Result/Conclusion] A medical case metadata model oriented toward TCM diagnosis and treatment knowledge bases was constructed, providing a comprehensive description of the TCM diagnosis and treatment process and offering a reference for standardizing TCM case metadata.

Keywords: metadata; metadata model; Traditional Chinese Medicine medical case; Traditional Chinese Medicine diagnosis and treatment knowledge base; Traditional Chinese Medicine knowledge organization

Classification Number: G250

DOI: 10.13266/j.issn.0252-3116.2021.02.001

Traditional Chinese Medicine represents a unique cultural and scientific heritage of China. As a vehicle for TCM inheritance, medical cases embody the valuable clinical experience and syndrome differentiation and treatment thinking of practitioners, serving as an essential pathway for scholars to learn from predecessors' academic ideas and clinical experience [1]. With the growth of medical case data and the deepening of modernization research in Chinese medicine, case information management has gradually shifted from traditional paper-based systems to modern information technology and database management. TCM diagnosis and treatment databases and knowledge bases are data and knowledge platforms capable of storing, sharing, mining, and analyzing TCM cases. Compared with treatment databases, knowledge bases place greater emphasis on knowledge organization and mining functions. Currently, China has established several TCM case information databases and knowledge bases, such as the "Traditional Chinese Medicine Basic Database System" [2], "Statistical Analysis System for

Historical Medical Cases” [3], and “Chinese Traditional Medicine Literature Database” [4]. However, these databases have experienced suboptimal usage rates and poor interoperability, leaving massive case resources underdeveloped and underutilized. The main problems include: (1) Most databases contain historical case data with personalized characteristics, resulting in non-uniform metadata formats; (2) Existing case databases or knowledge bases lack detailed metadata, failing to comprehensively and structurally reveal details of the TCM diagnosis and treatment process, which impedes high-precision retrieval and knowledge discovery. Addressing these issues, this paper aims to establish case metadata standards for TCM diagnosis and treatment knowledge bases, proposing a TCM case metadata model based on systematic research to standardize case information recording formats, form a foundation for deeply revealing case contents, and promote better development and utilization of TCM case knowledge resources.

1. Current Status of TCM Medical Case Metadata Research

1.1 Connotation and Development of TCM Medical Cases

Medical cases, also known as diagnostic records, pulse records, treatment plans, or patient records [5], are continuous documentation of a physician’s diagnosis and treatment of patients, including syndrome differentiation, therapeutic principles, and prescription medication [6]. Overall, TCM medical cases represent the most original and reliable records of the diagnosis and treatment process, preserving vast amounts of information about disease treatment and serving as the fundamental knowledge unit of TCM diagnosis and treatment knowledge bases.

After more than two thousand years of development, medical cases have evolved from initially recording only basic patient information, disease patterns, etiology, pathogenesis, treatment methods, efficacy, and prognosis [7] to modern “medical record-style” formats referencing Western medical records. These modern formats include basic patient information, visit time, chief complaints, present illness history, past medical history, TCM four diagnostic methods, auxiliary examination results, diagnosis, and treatment recommendations. However, medical cases and medical records are similar but not identical. Both are treatment records, but medical records comprehensively document patient health status and disease processes according to modern hospital models, while medical cases focus on recording TCM clinical facts and diagnostic thinking activities [8]. The differences are shown in Table 1 [9]. Medical records form the foundation of medical cases, documenting patient health status and treatment methods, while medical cases provide more detailed records than medical records, including the physician’s syndrome differentiation process. This differentiation process represents the externalization of tacit knowledge from clinical experience, mostly found in the commentary sections of medical cases.

Internationally, CDISC (Clinical Data Interchange Standards Consortium) has established a series of clinical research data standards for electronic acquisition, exchange, submission, and archiving of clinical data [13]. China's TCM clinical research case report forms are designed based on CDISC's CDASH (Clinical Data Acquisition Standards Harmonization) standards, with added TCM clinical data features, such as TCM diagnosis content, to promote data exchange and sharing in TCM clinical research and support integration with clinical workflows [14].

1.2 Existing Medical Standards and Specifications

Both Chinese and Western medicine have medical record standards, but medical case standards are rare. Since medical records form the basis of medical cases, this paper builds upon existing medical record standards, adding detailed metadata from the TCM diagnosis and treatment process to comprehensively express case contents.

1.2.1 TCM Metadata Standards As a major country in scientific data, China has accumulated substantial research and data in TCM studies [10]. The Chinese Academy of Medical Sciences and China Academy of Chinese Medical Sciences have jointly developed the “Medical and Health Scientific Data Sharing Metadata Standard” and “Traditional Chinese Medicine Scientific Data Sharing Metadata Standard,” providing references for TCM database construction and metadata structure design. Internationally, China has conducted standardization work for traditional medicine in ISO/TC 215 (Health Informatics Technical Committee) in recent years [11], having released three TCM information standards, including ISO/TS 17938 “Traditional Chinese Medicine Literature Metadata,” which provides a common set of descriptive elements for TCM literature resources [12].

1.2.2 TCM Medical Record and Clinical Data Standards Regarding TCM medical record standards, with the advancement of integrated Chinese and Western medicine, TCM diagnosis and treatment records have gradually transformed into “medical record-style” text preservation. China has issued a series of standards and specifications for TCM medical records, such as the “Basic Specifications for Writing TCM Medical Records” and “Guidelines for Writing TCM and Integrated Chinese-Western Medical Records.” These specifications include the general process and some elements of physician diagnosis and treatment. With rapid IT development, many paper-based records have been converted to electronic archives for better information management and preservation. China has promulgated standards for electronic medical records, such as the “Basic Dataset Standard for Electronic Medical Records” (WS445-2014). These standards include not only elements from paper records but also various data structures suitable for information preservation. As “Healthy China” has become an important national development strategy, the Ministry of Health released documents in 2009 regarding health records and health information, such

as the “Health Information Dataset Metadata Specification (WS/T305-2009).” To ensure comprehensive element sources, this study also referenced TCM diagnosis and treatment process-related books and templates, including “TCM Diagnostics,” “Dictionary of Diagnostics,” and “Traditional Chinese Medicine Apprenticeship Case Templates.”

1.3 Literature Review of TCM Medical Case Metadata

Currently, medical metadata research in China is still in its early stages. Using the search query “Subject = Traditional Chinese Medicine AND Keywords = Metadata” in CNKI with a deadline of April 23, 2020, yielded only 17 documents, most of which studied metadata for TCM medical case literature. Only four documents were related to TCM case metadata: two [15-16] analyzed the needs and research status of TCM case databases, pointing out that complex TCM cases require a comprehensive metadata scheme for unified description to establish databases with consistent data formats and structures; two [17-18] addressed metadata schemes and structural design for case databases, dividing case database structures into nine modules: basic case information, disease basic information, diagnosis, treatment, prescription, erroneous operations, contraindications, Western medicine-related information, and medical orders, with detailed refinement of these nine parts (see Figure 1 [Figure 1: see original paper]). However, these were only preliminary attempts with generalized structuring of treatment process knowledge and insufficient refinement of metadata elements.

In summary, research on TCM case metadata in China remains in its infancy, with preliminary academic exploration far from meeting the needs of TCM diagnosis and treatment knowledge organization and traditional cultural inheritance. This paper addresses these issues by refining the granularity of existing metadata based on medical record (or electronic medical record) standards and existing metadata standards, enabling more comprehensive and in-depth revelation of TCM diagnosis and treatment knowledge in knowledge bases and providing a foundation for further TCM knowledge retrieval and mining.

2. Principles and Process for Constructing TCM Medical Case Metadata Model

2.1 Metadata Model Construction Methods

Most metadata models are established by referencing existing metadata standards and reflecting specific business domains. For example, the artwork digitization metadata model was constructed by referencing existing standard systems and research frameworks, combined with researchers’ and managers’ metadata usage needs, modularizing metadata according to different functions in describing content and management mechanisms [19]. The digital museum metadata model adapted existing metadata standards through reclassification to create a dynamic model meeting resource requirements [20]. The geological dis-

ter emergency information metadata model was constructed based on practical emergency information resource classification [21]. Thus, metadata model construction generally employs one or a combination of two methods: extending or transforming existing standards, or designing models based on specific business processes. This paper's TCM case metadata model combines both methods, extending elements based on existing standards and optimizing elements through actual case verification to ensure accuracy, practicality, and meet requirements.

2.2 Metadata Extension Principles

Metadata extension principles mainly include two types [22]: (1) Horizontal extension, which adds a new metadata element without semantic repetition with existing elements; (2) Vertical extension, which imposes stricter semantic limitations on current metadata elements to generate new metadata qualifiers [23], enhancing descriptive precision. Since single-level element qualifiers cannot precisely describe elements, vertical extension may add qualifiers to qualifiers. In this paper, primary, secondary, and tertiary elements represent these levels.

2.3 TCM Medical Case Metadata Model Design Process

This study follows a three-step process of “preliminary element definition—element verification and optimization—metadata model establishment” (see Figure 2 [Figure 2: see original paper]) to determine TCM case metadata elements. First, through investigating TCM-related standards and existing metadata standards, the study understands the TCM diagnosis and treatment process to identify core and extended elements. Next, cases from websites are selected and analyzed using content analysis, compared with elements identified from standards to identify gaps, preliminarily defining metadata elements. Finally, through field investigation methods, the study observes and interviews at TCM diagnosis and treatment sites to optimize preliminarily defined elements, obtaining the final set of case metadata elements.

3. Preliminary Definition of TCM Medical Case Metadata Elements

3.1 Determining Core and Extended Elements

Referencing the aforementioned TCM medical case (record) standards and specifications, and based on the “Basic Specifications for Writing TCM Medical Records” and the American SOAP medication record format, case metadata is divided into six core categories: basic case information, basic patient information, diagnostic process, diagnostic assessment, treatment measures, and analysis summary. Among these, diagnostic process, diagnostic assessment, and treatment measures reference the SOAP format. SOAP is a medication record format recommended by the American College of Clinical Pharmacy, where S represents Subjective (patient complaints, medical history, past history, etc.), O represents Objective (clinical examination results), A represents Assessment

(clinical diagnosis and treatment process analysis/evaluation), corresponding to “diagnostic assessment” metadata, and P represents Plan (including treatment principles, medication guidance, and related recommendations), corresponding to “treatment measures” metadata. “Analysis summary” is established based on the “commentary” section in “Traditional Chinese Medicine Apprenticeship Case Templates,” primarily providing review and summary of referenced prescriptions or historical references and treatment processes to give readers more detailed understanding of cases.

Each core metadata category includes several core elements (represented as primary elements in this paper). The core elements of “basic case information” and “basic patient information” reference six elements from DC metadata [24] (Title, Creator, Date, Source, Publisher, Language), two elements from CISM6F [25] (city, province or state), “Health Record Common Data Element Standard,” and “Basic Dataset Standard for Electronic Medical Records” (WS445-2014). The core elements of “diagnostic process,” “diagnostic assessment,” and “treatment measures” come from WS445.1-2014 Basic Dataset for Electronic Medical Records—Part 1: Medical Record Summary and WS445.2-2014—Part 2: Outpatient (Emergency) Medical Records. The core elements of “analysis summary” come from the “Traditional Chinese Medicine Apprenticeship Case Template.” To ensure comprehensiveness and detail, some core elements can be extended, with secondary elements vertically extended under primary elements for detailed description, and if necessary, tertiary elements extended under secondary elements (Table 2 shows sources for secondary and tertiary elements).

3.2 Analysis and Supplement of TCM Medical Case Examples

3.2.1 Investigation of TCM Medical Case Examples After determining core and extended metadata through literature analysis, content analysis was applied to online TCM case examples to prevent metadata omissions. The analysis coded specific metadata elements contained in each case, compared them with elements obtained in Table 2, and identified valuable diagnostic and treatment information elements not covered. First, TCM case examples were collected through web surveys, with samples selected for analysis. Next, analysis categories and coding tables were designed based on elements in Table 2. Finally, online TCM case examples were analyzed according to the coding table to identify unclassifiable diagnostic and treatment information elements for metadata supplementation.

Samples were primarily selected from frequently updated TCM case websites, including TCM Chinese Medicine Network, TCM Family Heritage—Case Experience, Ai-Ai-Yi Website, and Good Doctor Online. These four websites contain numerous TCM diseases and treatment prescriptions. This study required complete case examples, excluding articles describing only methods or partial cases. Ultimately, 50 articles meeting requirements were selected from each website, published before October 20, 2019, selected sequentially from the first page backward, totaling 200 articles.

3.2.2 Design of Categories and Coding Tables Categories are classification items in content analysis, corresponding to core metadata in this study. Six categories were established (see Figure 3 [Figure 3: see original paper]): basic case information, basic patient information, diagnostic process, diagnostic assessment, treatment measures, and analysis summary.

Coding was conducted based on the aforementioned elements, with each code representing the final-level element of a core metadata category. For example, an extension under “diagnostic process” might be: “present illness history”—“palpation”—“pulse diagnosis,” where “pulse diagnosis” is the final-level element of “diagnostic process,” coded as “pulse diagnosis.” The coding design consisted of two parts: (1) pilot coding to test inter-coder reliability, and (2) formal content coding.

- (1) **Pilot Coding for Reliability Testing:** Reliability is defined as consistency among different coders in content classification. To ensure accurate coding, inter-coder reliability must be assessed. This study selected two coders: the author and a research group member with similar research interests. Before pilot coding, the other coder was trained on element meanings and corresponding sentences. Reliability testing standards generally require an overall inter-coder agreement level above 85% within a 95% confidence interval to pass [26].

First, the sample size for reliability testing was determined. According to the theory on required sample sizes for reliability testing with different population sizes proposed in “Content Analysis: Quantitative Research of Media Information” [27], with 200 total case samples and assuming 90% overall coder agreement within a 95% confidence interval, formula (1) indicates that 68 samples needed to be randomly selected for reliability testing.

$$(N-1)(0.03)^2 + PQN / (N-1)(0.03)^2 + PQ \text{ (Formula 1)}$$

Where n = reliability testing sample size, N = total study sample size, P = overall agreement level, $Q = 1-P$.

Next, two coders coded 68 case examples. Uncertain coding situations were recorded and analyzed by a third party (a TCM student) to determine coding assignments. Results were: $N_1=93$, $N_2=91$, $M=85$, $n=2$. Substituting into the reliability formula (Formula 2):

$$R = nK / 1 + (n-1) \times K / N_1 + N_2 \text{ (Formula 2)}$$

Where n = number of coders, M = number of elements both coders completely agreed on, N_1 = number of codes by first coder, N_2 = number of codes by second coder. K = inter-coder agreement.

The inter-coder agreement was $K = 92\%$, with reliability $R = 96\%$, exceeding the 90% threshold and passing reliability testing, allowing continued coding.

- (2) **Formal Content Coding:** After passing reliability testing, remaining case examples were coded. As with pilot coding, uncertain situations

were recorded and resolved with a third party (TCM student), resulting in the coding table shown in Table 3 .

3.3 Reliability Testing

Reliability refers to the accuracy and dependability of test results [28], generally judged by the degree of difference between measured and evaluated values. To verify reliability of elements obtained through content analysis, two other research group members served as evaluators for metadata element analysis. The coding table was provided to both evaluators, who received training on coding methods and element meanings with examples, then practiced coding 5 identical cases to verify understanding. After confirming their mastery, 40 additional cases from the four websites were jointly coded for reliability testing. Substituting results into Formula (2) yielded inter-coder agreement $K = 0.85$ and reliability $R = 0.92$, meeting acceptable reliability standards and confirming the coding table's usability.

3.4 Determining Supplementary Metadata Elements

Through content analysis of case examples, a coding table was obtained (see Table 3). Comparison with elements in Table 2 revealed newly supplemented elements (jointly determined by both coders as belonging to certain core metadata categories, shown in Table 4 with gray-shaded font indicating new elements).

One element was not found in web survey cases: “diagnostic process—present illness history—inspection—inspection of infant finger venule.” Since further optimization would occur through field investigation, this element was temporarily retained for potential deletion after field investigation.

4. Optimization of TCM Medical Case Metadata Elements and Model Establishment

4.1 Element Verification and Optimization Process

Based on literature analysis and web case content analysis alone, the obtained elements could not be fully confirmed as matching actual diagnosis and treatment processes. Field investigation methods were required for verification and optimization. Field investigation is a fundamental approach for researchers to obtain data, emphasizing personal observation and interviews in real social contexts, with maximum advantages in directness and reliability. Forms include direct observation, interviews, experiments, etc. [29]. However, parts of TCM diagnosis are tacit and inexpressible, with experienced physicians' diagnostic processes not fully articulated. Researchers requested physician clarification on questionable observations to avoid missing element selection.

This study employed direct observation and interview methods for further optimization. The process included three phases: preparation, field investigation, and data organization.

- (1) **Preparation Phase:** Before site visits, elements obtained from standards and case examples were organized into table format and printed as element sheets for recording. Recorders were the two coders (the author and a research group member).
- (2) **Field Investigation Phase:** Ten TCM physicians in Changchun and Songyuan were selected. Before investigation, physicians were informed that the purpose was to understand their diagnostic processes, with researchers observing and recording behaviors and patient conversations related to conditions without involving patient privacy. Printed element sheets were provided for physicians to review and comment on element rationality and practicality.

During observation, both coders recorded processes using element sheets, marking questionable areas for discussion during breaks. In breaks, researchers organized questions, resolved internally solvable issues, unified questions requiring physician responses, and conducted interviews. Interview questions and answers are shown in Table 5 .

- (3) **Data Organization Phase:** During the first physician interview, the physician noted that elements in “basic case information” and “basic patient information” were automatically generated by the medical record system and asked the author to verify this.

After each physician investigation, both coders organized recorded content, counted element occurrences, and unified disputed parts to determine supplementary elements. A total of 972 case records were obtained, with elements organized as shown in Table 6 (presented as final-level elements of core meta-data; gray-shaded areas indicate new elements).

4.2 Element Optimization Results

The element “diagnostic process—present illness history—inspection—inspection of infant finger venule” did not appear in either case investigation or field survey results. After discussion among researchers, it was decided to delete “inspection of infant finger venule,” with provision for future addition if needed during data entry.

Compared with elements from case content analysis (Table 3), field investigation yielded additions and deletions. Added elements (Table 7) and reasons are as follows:

- (1) **Stethoscope:** A common tool for physicians to examine internal organs. Since one field investigation physician was a pediatric specialist who frequently used stethoscopes for heart and lung examination in over 90% of pediatric cases, “stethoscope” was vertically extended as a secondary element under primary element “physical examination.”
- (2) **Percussion sound:** “Auscultation” in TCM diagnostics includes listening

to sounds and smelling odors. Cases generally show “listening to sounds” as hearing patient breathing, coughing, sighing, etc. However, field investigation found some physicians judge conditions by tapping body parts (e.g., abdomen) and listening to sounds. Researchers defined this as “percussion sound,” extending it as a tertiary element under secondary element “auscultation.”

- (3) **Inquiry about height and weight:** Generally unrelated to conditions and not expressed in cases. However, TCM physicians consider height and weight for medication dosage, requiring vertical extension as a tertiary element under secondary element “inquiry.”
- (4) **Inquiry about mental state:** Refers to asking about patients’ mental status after illness onset. During field investigation, patients reported post-illness irritability and poor sleep quality unrelated to daily life stress. This distinction requires separate representation as a tertiary element “inquiry about mental state” under secondary element “inquiry,” distinct from “mental status” in “personal history.”
- (5) **Trauma history:** Refers to verifiable past injuries. Field investigation revealed patients with histories of serious car accidents or fractures. Referencing “Admission Record Writing Requirements and Format,” “surgical trauma history” was split from “past history,” with “trauma history” vertically extended as a secondary element under primary element “past history.”

4.3 Metadata Model Establishment

Through the aforementioned optimization process, the final TCM medical case metadata model was established (see Figure 4 [Figure 4: see original paper]). This model covers the complete TCM diagnosis and treatment process, clarifies case content formats, reduces barriers to case information sharing and acquisition, and facilitates case information preservation, management, and retrieval.

This metadata model reveals key content in the TCM diagnosis and treatment process more comprehensively and deeply than existing TCM case standards and specifications. The TCM diagnostic process contains substantial personal experience (tacit knowledge) rarely shown in regular medical records, potentially leading to incomplete knowledge mining. This paper constructs a scientifically validated TCM case metadata model by referencing existing standards and models, analyzing actual cases, and conducting field investigations, aiming to increase the externalization of physicians’ tacit knowledge and enable better development and utilization of case resources.

As a precious cultural heritage, TCM and its medical cases are crucial for inheritance and development. However, current online case information sharing is hindered and retrieval is imprecise, making unified formats a primary consideration. This paper adopts a theory-practice integration approach, first referencing

TCM standards and metadata specifications to obtain core elements and extensions; second, collecting TCM case examples to study content characteristics and preliminarily define elements through content analysis; third, optimizing elements through field observation and interviews; and finally establishing the metadata model. This high-precision model enables deep structural organization of treatment process information, facilitating multi-dimensional retrieval and knowledge mining in diagnosis and treatment databases or knowledge bases, and deepening theoretical research and practical work in TCM knowledge organization.

References

- [1] Cui Junbo, Chen Baogui. Discussion on Inheritance Ideas and Experience of Famous TCM Master Inheritance Studios[J]. Medical Information, 2015(10): 273-273.
- [2] Ren Tingge, Liu Xiaofeng, Gao Jianbo, et al. Introduction to the “Traditional Chinese Medicine Basic Database System”[J]. Chinese Journal of Traditional Chinese Medicine Information, 2001, 8(11): 90-91.
- [3] Wu Zhongping, He Xinhui, Ke Xuefan, et al. Introduction to the Medical Case Query and Statistical Analysis System[J]. Shanghai Journal of Traditional Chinese Medicine, 2003, 37(3): 54-56.
- [4] Yu Qi, Cui Meng, Li Yuanbai, et al. Research on Construction Standards for Traditional Chinese Medicine Literature Databases[J]. World Science and Technology-Modernization of Traditional Chinese Medicine, 2014, 16(11): 2304-2307.
- [5] Su Li. Introduction to Traditional Chinese Medicine Case Studies[M]. Beijing: People’s Medical Publishing House, 2009.
- [6] China Academy of Traditional Chinese Medicine. Selected Explanations of Traditional Chinese Medicine Terms[M]. Beijing: People’s Medical Publishing House, 1973.
- [7] Gao Dianpu, Wang Yinghui, Zhang Runshun, et al. Review of Standardization Research on Traditional Chinese Medicine Cases[J]. Journal of Traditional Chinese Medicine University of China (Clinical Medicine), 2013, 20(5): 3-6.
- [8] Sun Guangrong. Ideas and Methods for Medical Case Research and Writing[J]. Journal of Beijing University of Traditional Chinese Medicine, 2011, 26(11): 2369-2373.
- [9] Yang Jianyu, Li Yanzhi, Fan Zhuwen, et al. Discussion on Medical Case Writing and Editing—Concurrently Discussing Research on Academic Thought of Master Sun Guangrong’s Typical Case Studies[J]. Guangming Traditional Chinese Medicine, 2011, 26(11): 2369-2373.
- [10] Li Jibin, Zou Xiaonong, Qiao Youlin. Metadata and Its Application in Medical Scientific Data Sharing[J]. Medical Intelligence Work, 2006, 27(1): 10-14.
- [11] Li Haiyan, Cui Meng, Ren Guanhua, et al. Progress in ISO/TC 215 Traditional Medicine Information Standardization Work[J]. International Journal of Traditional Chinese Medicine, 2011, 33(3): 193-195.

- [12] China Traditional Chinese Medicine News. ISO Releases First Two Traditional Chinese Medicine Information International Standards—“Traditional Chinese Medicine Language System Semantic Network Framework” and “Traditional Chinese Medicine Literature Metadata”[J]. *Jiangsu Traditional Chinese Medicine*, 2014, 46(10): 27-27.
- [13] Clinical Trial Electronic Data Capture Task Group. Electronic Clinical Data Capture Position Paper Revision 1[EB/OL]. [2020-05-14]. <https://rammellconsulting.files.wordpress.com/2011/05/edc-position-paper-revision-1-final-version.pdf>.
- [14] Li Geng, Li Xiaoyan, Wen Zehuai. Application of CDISC Standards in Traditional Chinese Medicine Clinical Research Case Report Form Design[J]. *Journal of Guangzhou University of Traditional Chinese Medicine*, 2014, 31(1): 138-141.
- [15] Wu Ye, Ma Lu. Current Status and Prospects of Traditional Chinese Medicine Case Database Development[J]. *Chinese Journal of Medical Library and Information Science*, 2012, 21(11): 49-52.
- [16] Song Yangwen, Tian Rui, Chen Jin, et al. Research on Utilization and Demand Survey of Traditional Chinese Medicine Case Databases[J]. *Journal of Medical Informatics*, 2013, 34(7): 63-66, 89.
- [17] Tian Rui, Ma Lu. Design of Metadata Scheme for Traditional Chinese Medicine Case Database[J]. *Chinese Journal of Medical Library and Information Science*, 2014, 23(10): 66-69.
- [18] Liu Baojie, Chen Jin, Ma Lu. Practice in Metadata Structure Design for Traditional Chinese Medicine Case Database[J]. *Journal of Medical Informatics*, 2013, 34(12): 70-73.
- [19] Weng Ziyang, Li Zhuoran. Construction of Artwork Digitization Metadata Model[J]. *Digital Library Forum*, 2019(6): 11-18.
- [20] Jia Junzhi, Shi Xuan. Research on Digital Museum Metadata Standard Construction[J]. *Journal of Shanxi University (Philosophy and Social Science Edition)*, 2015, 38(1): 114-119.
- [21] Li Li. Research on Geological Disaster Emergency Information Resource Metadata Model Construction[D]. Nanchang: Nanchang University, 2014.
- [22] Zhu Yunqiang, Xiang Wenxin, Feng Min, et al. Research on Extensible Environmental Protection Archives Metadata[J]. *Archives Science Bulletin*, 2012(5): 78-82.
- [23] Yi Junkai, Zhou Yubin, Chen Gang. Research and Practice on Extensible Digital Museum Metadata Specifications[J]. *Digital Library Forum*, 2014(2): 43-53.
- [24] DCMI Usage Board. Dublin Core Metadata Element Set, Version 1.1: Reference Description[EB/OL]. [2020-05-14]. <https://www.dublincore.org/specifications/dublin-core/dces/>.
- [25] CISMéF Team. CISMéF: Catalog and Index of French-language Health Internet Resources, Quality-controlled Subject Gateway[EB/OL]. [2020-05-10]. <http://www.chu-rouen.fr/cismef/cismefeng.html>.
- [26] Ding Xiaojian. Research on Content Analysis of Aviation Enterprise Weibo and Its Impact on User Attention Behavior[D]. Shanghai: Donghua University,

2014.

[27] Riffe, Lacy, Fico, et al. Content Analysis: An Introduction to Its Methodology[M]. Translated by Ji Meiyun. 2nd ed. Beijing: Tsinghua University Press, 2010.

[28] Shi Ronghua. Dictionary of Social Psychology[M]. Chengdu: Sichuan People's Publishing House, 1988.

[29] Hu Jingqing. Field Investigation Should Become One of the Basic Methods for TCM Basic Theory Research[J]. Chinese Journal of Basic Medicine in Traditional Chinese Medicine, 2014, 20(12): 6-8.

Author Contributions:

Li Lu: Data investigation and organization, paper writing and revision;

Ma Jie: Topic selection, research framework design, paper revision and finalization;

Sun Hengyu: TCM knowledge explanation, analysis result consultation, paper revision;

Wang Jue: Data collection and analysis.

The Construction of Medical Case Metadata Model for the Knowledge Base of Traditional Chinese Medicine Diagnosis and Treatment

Li Lu¹, Ma Jie^{1,2}, Sun Hengyu³, Wang Jue¹

¹ School of Management, Jilin University, Changchun 130022

² Information Resources Research Center, Jilin University, Changchun 130022

³ The First Clinical Hospital of Jilin Academy of Traditional Chinese Medicine, Changchun 130022

Abstract: [Purpose/significance] In the process of traditional Chinese medicine (TCM) medical record management, problems such as non-uniform metadata and insufficient precision hinder information retrieval, knowledge sharing, and mining in TCM medical record knowledge bases. Establishing a standardized TCM medical record metadata model plays an important role in supporting deep multi-dimensional retrieval, knowledge sharing, and mining of TCM medical records. [Method/process] This paper summarizes existing TCM-related standards and metadata standards through literature survey; collects TCM medical record examples through network survey, analyzes medical record content through content analysis, and preliminarily defines medical record metadata elements; then verifies and optimizes the preliminarily defined medical record metadata elements through field investigation method by observing and visiting the TCM diagnosis and treatment process on site, and obtains the final set of medical record metadata elements. [Result/conclusion] A medical record metadata model oriented to the TCM diagnosis and treatment knowledge base is constructed, which comprehensively describes the TCM diagnosis and treatment process and provides a reference for unifying TCM medical record metadata.

Keywords: metadata; metadata model; traditional Chinese medicine medical record; traditional Chinese medicine diagnosis and treatment knowledge base;

traditional Chinese medicine knowledge organization

2018-2020 Library and Information Service Excellent Papers

Since 2014, this journal has published the TOP 10 highly cited papers of the current year and the previous two years. In 2020, Library and Information Service will continue the 2019 selection method, selecting highly cited and highly downloaded papers published in the past 3 years, based on annual TOP 50 and considering publication time, with the editorial office finally selecting 10 papers for announcement. The 2018-2020 Library and Information Service excellent papers are:

| Title | Authors | Publication Year/Issue |
|--|-------------------------------------|------------------------|
| Review of Domestic Online Health Community Research Status | Liao Haihan, Wang Yuefen, Guan Peng | 2018, 62(9) |
| Blockchain Technology and Its Transformative Impact on Library Development | Cao Shujin, Liu Huiyun | 2018, 62(13) |
| Topic Mining and Opinion Identification of Different Disseminators in Weibo Public Opinion Communication Cycle | Shao Bo, Zhang Wenzhu | 2018, 62(19) |

| Title | Authors | Publication Year/Issue |
|--|--|------------------------|
| Library Transformation from Resource Capability to Service Capability | Chu Jingli, Zhao Yan | 2019, 63(1) |
| Reader-Centered Smart Library Research Practice and Thinking on Next-Generation Library System Platforms | Wang Bing, Wu Chao | 2019, 63(1) |
| The Origin, Evolution Trend and Meaning of Security Intelligence Concept—Speculation from the Perspective of Safety Science Theory | Zhou Tao, Tan Qi, Takirova Bayan, Deng Shengli | 2019, 63(3) |
| Research on the Mechanism of Social Interaction's Effect on Users' Knowledge Payment Intention | Mou Dongmei, Shao Qi, Han Nannan, Wang Ping, Jin Shan, Jin Chunyan | 2019, 63(4) |
| | Wang Xiwei, Liu Weili, Jia Fengqi, Zhang Chuang | 2020, 64(3) |

| Title | Authors | Publication Year/Issue |
|--|---------|------------------------|
| Multi-dimensional Social Attribute Analysis and Visualization of Weibo Public Opinion—A Case Study of a Vaccine Incident | | 2020, 64(5) |
| Influencing Factor Model and Empirical Research on Live Streaming APP Usage Behavior | | 2020, 64(5) |

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.