

Visual Analysis of North-South Clinical Treatment Differences Based on the "Disease-Physician-Prescription-Herb" Association

Chenyuan Wu¹, Chao Chen¹, Liuying Li^{2,*}, Xianjun Ning¹, Qian Nie¹, Tiancheng Yang¹,

Rongju He¹, Ying Guo¹

¹ Sichuan University of Science & Engineering, Yibin Sichuan, 644000, China

² Zigong First People's Hospital, Zigong, Sichuan, 643000, China

* Corresponding author: Liuying Li

Abstract: As a pivotal component of China's Intangible Cultural Heritage, Traditional Chinese Medicine (TCM) diagnosis and treatment holds unique value in the prevention and treatment of complex diseases and in health management, demonstrating significant North-South regional disparities. To deeply unravel the intrinsic patterns underlying these disparities, this paper constructs a visual analytics system based on medical records, prescriptions, and medicinal material data. First, data is categorized according to the characteristics of TCM diagnosis and treatment, and visualization tasks are defined. Then, by integrating text, spatial, and multidimensional visualization techniques, along with a multi-view linked interaction mechanism, the system enables the correlation analysis of textual features, geographical distribution, and prescription formulation patterns in north-south medical records. Comparative examination of prescriptions and herb combinations utilized by northern and southern practitioners reveals substantive divergences in TCM diagnostic, therapeutic strategies, and pharmacopeia applications. This research contributes both empirical foundations for standardizing TCM clinical practice and pharmaceutical protocols, and a new method for the digital research of TCM cultural heritage.

Keywords: TCM Culture; Visual Analytics; Prescription; Herb.

1. Introduction

Traditional Chinese Medicine (TCM) diagnosis and treatment, as a core component of China's intangible cultural heritage, demonstrate unique advantages in preventing and treating complex diseases. Its diagnostic framework, grounded in the principles of "holism" and "treatment based on syndrome differentiation" [1], This cognitive model, emphasizing the cor-responce between humans and nature, regulates the dynamic balance between the internal and external environments of the human body, forming a diagnostic and therapeutic framework complementary to modern medicine. It not only profoundly influences the health practices of the Chinese population but also provides Eastern wisdom for the development of global medicine. Notably, TCM practices exhibit significant regional heterogeneity. Research by Huang [2] indicates that physicians from northern and southern China differ in their focus on syndrome differentiation, particularly in distinguishing patterns of cold, heat, deficiency, and excess. Such differences manifest in various aspects, including the number of medicinal ingredients and dosage adjustment strategies in formula composition. Moreover, the geographical characteristics of medicinal materials further reinforce the regional distinctiveness of TCM diagnosis and treatment, contributing to the rich diversity of clinical practices.

However, current visual analytics research addressing North-South regional disparities in TCM remains insufficient. Traditional literature analysis methodologies struggle to effectively represent the complex network associations among symptoms, prescriptions, and herbs, often leaving TCM knowledge in the form of "tacit experience." In the

context of TCM research's transition toward a data-driven paradigm, visualization technology offers a robust methodology for knowledge representation and the discovery of underlying diagnostic patterns. This paper presents a visual analytics system designed to explore regional diagnostic cultures. The system facilitates a comparative analysis of how renowned practitioners from different regions address identical syndromes, specifically examining herbal compositions and the spatial distribution of both physicians and medicinal resources. By integrating extensive unstructured medical records and employing a multi-view linkage framework, the system overcomes the limitations of traditional research, transforming multi-source text into interactive visual representations. This study aims to enhance the understanding of TCM's regional nuances while providing new methodologies for both standardization and personalized medicine. By visually manifesting diagnostic disparities, this work seeks to revitalize TCM innovation and promote its integration into modern healthcare systems.

2. Related Work

Existing research on the integration of TCM diagnostic culture and visualization includes TCM syndrome visualization [3], network correlation of TCM compound prescription medication patterns [4], and TCM knowledge graphs [5]. By integrating factors such as prescriptions, syndromes, diseases, medication patterns, and medicinal materials, these studies utilize big data visual analysis [6] and knowledge graphs [7] to present TCM research achievements and promote the revitalized inheritance of TCM. However, few studies have focused on the regional differences between Northern and Southern TCM. This paper, based on TCM

diagnostic records, primarily utilizes textual data supplemented by spatial and multi-dimensional data. It combines spatial and multi-dimensional visualization techniques to conduct a differential analysis of TCM diagnosis between the North and the South.

2.1. Spatial Data Visualization

Spatial visualization employs technologies such as computer graphics and Geographic Information Systems (GIS) [8] to convert abstract spatial data, encompassing geographic coordinates, distribution patterns, relational networks, and spatiotemporal dynamics, into intuitive graphical forms including thematic maps, schematic diagrams, and three-dimensional models. By integrating interaction and analytical functionalities, it facilitates the understanding of spatial patterns and supports decision-making processes. Based on cognitive science and cartographic theory, Gao [9] systematically investigated the representation, analysis, and quality inspection of geospatial data, demonstrating the logical transition from abstraction to concrete visualization and its cross-disciplinary value. Wang [10] further advanced the application of spatial visualization by elevating maps from traditional information carriers to knowledge transmission tools. By merging geospatial visualization with knowledge visualization methodologies through ontology-driven approaches and data models, Wang achieved dynamic and multi-layered expressions of spatial knowledge to satisfy diverse cognitive requirements. This paper focuses on presenting the geographic distribution of renowned TCM practitioners and medicinal herbs, aiming to explore the spatial correlations and regional characteristics inherent in their medication patterns.

2.2. Multi-dimensional Data Visualization

Multi-dimensional data refers to variables that possess multiple attributes. Multi-dimensional data visualization is the process of transforming high-dimensional data into graphical or image-based representations that are easily interpretable by humans [11]. Ma et al. [12] utilized multi-dimensional visualization to perform correlation analyses on the temporal formation, vocal style fusion, and migration culture of Sichuan Opera, depicting its historical evolution. Hu [13] integrated Theme-River with time-series and scatter plots to characterize the stylistic evolution of Song poetry and the historical development of different literary schools.

To address the multi-dimensional information embedded in TCM medical records, this paper links Northern and Southern practitioners through syndromes and integrates the geographic information of both renowned physicians and medicinal herbs. By combining the composition of prescriptions and herbal ingredients, the proposed system facilitates an intuitive investigation of the similarities and differences in regional formulas, as well as the distributional disparities of medicinal materials and practitioners. Ultimately, this work constructs a multi-dimensional linkage framework spanning from syndromes and formulas to herbs and geographic regions.

3. Data and Tasks

3.1. Task Analysis

TCM diagnostic knowledge, especially the expertise of renowned veteran practitioners with distinct regional characteristics, primarily exists as highly personalized and

hard-to-formalize tacit knowledge. For centuries, its transmission has heavily relied on the master-apprentice model, characterized by oral instruction and prolonged experiential learning. This reliance results in a knowledge system that is highly fragmented and subjective, posing significant challenges for systematic integration and scalable application. Although medical records are the primary repositories for capturing such tacit knowledge, conventional literature-based research methods, limited by manual induction and individual case studies, are inadequate for processing massive volumes of unstructured data. Specifically, these methods struggle to effectively extract the complex, non-linear associations among symptoms, prescriptions, and herbs. Consequently, the understanding of regional diagnostic styles remains largely empirical, lacking the support of verifiable and objective patterns. To address these challenges, this paper defines the following analytical tasks:

1. Entity Recognition in TCM Medical Records

TCM medical records contain complex multidimensional data, including practitioners, herbs, prescriptions, symptoms, and dosages. It is therefore necessary to identify and classify entities within these records to generate structured data. This provides the foundational data support for visual design and facilitates further data collection and processing.

2. Reconstruction of Medical Entities

Focusing on core entities such as diseases, medicinal herbs, and prescriptions, this task involves integrating professional medical literature to establish standardized and structured frameworks for herbal information, prescription compositions, and disease retrieval.

3. Visual Presentation

The system aims to design a visualization interface that conforms to TCM cognitive characteristics, enabling users to perform exploratory analysis on multidimensional data including herbs, regions, prescriptions, and symptoms. By leveraging multi-view linkage, the system achieves the intuitive display and dynamic querying of complex relationships, thereby facilitating the understanding and dissemination of diagnostic knowledge.

3.2. Data Sources

Based on a systematic task analysis framework, this study established a data acquisition pipeline to support the analysis of North-South regional diagnostic variations in TCM. The data corpus primarily comprises clinical records from the "Collection of National Famous TCM Medical Records" [14] and the "Essence of Modern Famous Chinese Medicine Case Records (Volumes 1-6)" [15]. Prescription composition data were cross-referenced with "Formulas of TCM" [16], while herbal origins and quality standards were verified against "TCM Materia Medica" [17] and the "Pharmacopoeia of the People's Republic of China" [18]. Biographical details of renowned practitioners were retrieved from the "Famous Doctors Encyclopedia" website. By integrating these multi-source heterogeneous data, a structured database encompassing diagnostic decision-making, formula compatibility, and herbal properties was constructed, providing the empirical foundation for subsequent feature mining and comparative analysis.

3.3. Data Classification

In the field of visual analytics, data classification involves organizing raw data into distinct categories or dimensions

based on their attributes, characteristics, or structures. This process enables the selection of appropriate visualization techniques for representation, correlation, and analysis. Given the complexity of medical record data within the context of North-South regional disparities, this study refined and

categorized information regarding TCM formulas, medicinal herbs, renowned practitioners, and other auxiliary metrics extracted from medical texts. The specific classification results are presented in Table 1.

Table 1. The specific classification results

Data Classification	Contents
TCM Herbs	Therapeutic effects, origin, herb name, herb image
TCM Formulas	Formula name, formula alias, herbal composition, formula efficacy
Renowned Practitioners	Practitioner profiles, geographical location, profile images
Others	Symptoms

4. Data Processing

This section describes the data processing workflow and methodology. The framework of the processing workflow and system design is illustrated in Figure 1. During the data collection phase, to achieve the research goal of analyzing diagnostic differences between Northern and Southern TCM, the system integrates multi-source heterogeneous data. This primarily includes clinical records, classical texts on prescriptions and medicinal herbs, and authoritative online encyclopedias of renowned TCM practitioners.

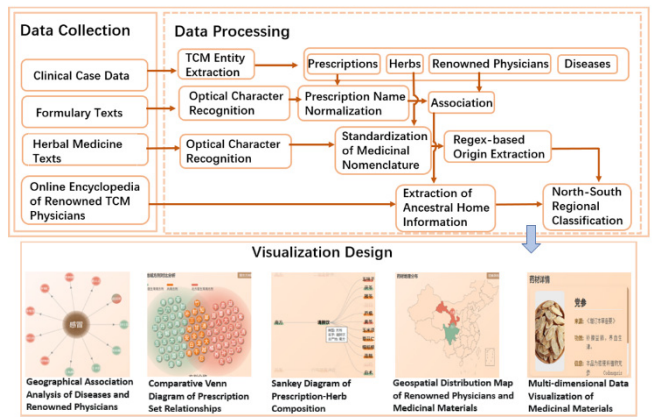


Figure 1. Data Processing Flowchart

In the data preprocessing stage, a systematic structured workflow was developed for the multi-source TCM data. First, clinical records serve as the core analytical objects. Unstructured text is automatically parsed using Named Entity Recognition to extract key entities such as prescriptions, herbs, symptoms, and practitioners, providing a structured foundation for subsequent analysis. Based on the practitioners identified from the records, their geographical origins are retrieved from online encyclopedias to categorize the practitioner groups into Northern or Southern regions. For prescription entities extracted from medical cases, entity normalization is performed by referencing standardized prescription names digitized from classics such as "Fangji Xue" (Science of Prescriptions) using OCR technology. These normalized prescriptions are then linked to the practitioners and their regional information to achieve the North-South classification of prescription data. Finally, after standardizing the nomenclature of herb data obtained via OCR, regular expression matching is employed to extract origin descriptions from the text. This establishes the North-South regional attributes for herbs, forming a complete geographical classification system for medicinal materials.

In the visual design phase, the system leverages D3.js and other front-end technologies to construct an interactive visual

analytics interface featuring five coordinated views. Through a multi-view linkage mechanism, the system reveals the medication patterns and regional characteristics of Northern and Southern TCM across multiple dimensions, including spatio-temporal distribution, association, and comparison.

5. System Design

Based on the task analysis, this paper designs and implements a visual analytics system centered on the "Disease-Syndrome-Prescription-Herb" association to explore regional differences in TCM clinical practices. The system interface is shown in Figure 2. The framework adopts (B) the Geographical Association Analysis of Diseases and Renowned Physicians as its visual centerpiece, complemented by a suite of coordinated views: (C) the Comparative Venn Diagram of Prescription Set Relationships, (D) the Sankey Diagram of Prescription-Herb Composition, (E) the Multi-dimensional Data Visualization of Medicinal Materials, and (F) the Geospatial Distribution Map of Renowned Physicians and Medicinal Materials. Together, these components form a multi-level visual analytics framework.

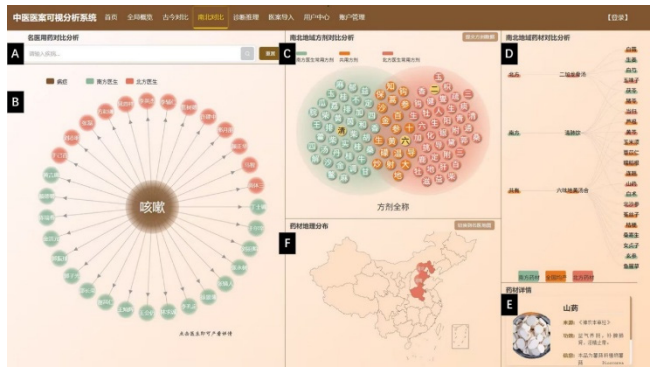


Figure 2. Visual Analytics System

5.1. Geographical Association Analysis of Diseases and Renowned Physicians

After entering a symptom name in the search bar (A), the system dynamically generates the Geographical Association Analysis of Diseases and Renowned Physicians (B). This view features the target symptom as a brown central node, connected radially to Northern and Southern practitioners. Node colors are automatically encoded based on entity types and regional attributes: Southern practitioners are represented in cyan-green, Northern practitioners in pale vermillion, and those common to both regions use a cyan-red gradient, supported by a legend for rapid identification. This design allows users to explore the relationships between regional practitioners for a specific disease in depth. Additionally, this

view is linked with the Geospatial Distribution Map of Renowned Physicians and Medicinal Materials (F); by using a toggle button, users can intuitively visualize the geographic distribution of practitioners associated with the selected symptom.

5.2. Comparative Venn Diagram of Prescription Set Relationships

To deeply analyze the differences and commonalities in prescription selection between Northern and Southern TCM schools, the system features a Venn-based regional prescription comparison view (C). This view provides an intuitive representation of the set relationships between Southern prescriptions, Northern prescriptions, and their intersection. It effectively characterizes the geographical distribution of prescriptions, providing visual support for the comparative analysis of regional medication strategies.

Based on the geographical origins of the practitioners recorded in the medical cases, the system automatically categorizes prescriptions as either "Southern" or "Northern" and visualizes them using a dual-circle overlapping Venn diagram. The left domain represents the Southern prescription set, the right domain represents the Northern prescription set, and the overlapping area indicates prescriptions commonly used by both Northern and Southern practitioners. Each node within the view corresponds to a specific prescription. This visualization allows users to rapidly identify both medication commonalities and regional specificities across different schools. Furthermore, the system provides a "Submit Prescription Data" interactive control, enabling researchers to select specific subsets of prescriptions to trigger a linked update of the Sankey Diagram (D) in the downstream analysis module. This interaction extends the analytical chain and deepens the exploration of prescription-herb relationships.

This module offers significant practical value for preserving regional medication characteristics and constructing a localized framework for TCM diagnostic knowledge.

5.3. Sankey Diagram of Prescription-Herb Composition

To further analyze the selection differences and commonalities of medicinal herbs between Northern and Southern prescriptions, the system automatically generates a Sankey diagram (D) once the user selects a specific combination of prescriptions. This diagram constructs a hierarchical medication path from geography to prescriptions and finally to constituent herbs, serving as the core visual component of the "North-South Regional Herbal Comparative Analysis" module. The left side of the Sankey diagram represents the regional source, which is categorized into "South," "North," and "Common" to reflect the geographical origin of the practitioners. The middle section displays the specific prescriptions selected by the user (in this example: *Huluba* from the South, *Yigankang* from the North, and the shared *Daochi San*). The right side lists the specific TCM herbs contained within these prescriptions, arranged according to their affiliations. Node colors are encoded to represent regional attributes: pale vermillion signifies Northern-exclusive herbs, cyan-green signifies Southern-exclusive herbs, and orange represents herbs common to both regions.

This view provides an interactive analytical platform for the teaching and dissemination of regional medical styles. It

holds significant methodological importance for promoting visual knowledge mining and cross-regional comparative research in TCM.

5.4. Multi-dimensional Data Visualization of Medicinal Materials and Geospatial Distribution Map of Renowned Physicians and Medicinal Materials

To enhance the system's knowledge depth and interactive coherence, we provide comprehensive functions for displaying herbal details and geospatial distributions. Selecting any herb node in the Sankey diagram (D) triggers an interaction linkage that presents the detailed attributes and spatial origin of that herb in the Multi-dimensional Data Visualization of Medicinal Materials (E) and the Geospatial Distribution Map (F), respectively.

The Multi-dimensional Data Visualization (E) dynamically displays the selected herb's name, literature source, and primary therapeutic effects. Sourced from authoritative TCM literature, the system employs a structured layout and key information highlighting to improve the efficiency of information retrieval for users. Simultaneously, the Geospatial Distribution Map (F) is activated. Built upon Chinese administrative divisions, this view highlights the primary production provinces of the herb and categorizes them into Northern and Southern regions using differentiated color encoding.

6. Case Analysis-The Impact of Regional Climate on Prescription Selection and Therapeutic Styles

The influence of regional climate on prescription selection and treatment styles is clearly observable within the diagnostic framework for "cough," as shown in Figure 3. This divergence is essentially a concrete manifestation of the TCM principle of "adapting treatment to local conditions," which systematically determines clinical treatment strategies and the principles of prescription compatibility.



Figure 3. Comparative Venn Diagram of Prescription Set Relationships

The Southern prescriptions concentrated in the pale green region are exemplified by "*Qingfei Yin*" (Lung-Clearing Decoction). Its core therapeutic methodology involves clearing, moistening, dispersing, and protecting Yin fluids, which precisely matches the humid and hot climate of the Southern region and the physical constitution prone to Yin deficiency and exuberant fire. The prescription utilizes

Scutellaria baicalensis (*Huangqin*) and Forsythia suspensa (*Lianqiao*) to clear heat and purge fire, while Poria cocos (*Fuling*) promotes diuresis to drain dampness, directly neutralizing damp-heat pathogens. As shown in Figure 4, the inclusion of Schisandra chinensis (*Wuweizi*) to astringe the lungs and nourish the kidneys embodies the Southern medical philosophy that "preserving one part of body fluid ensures one part of vitality." Simultaneously, it is paired with Atractylodes macrocephala (*Baizhu*) to strengthen the spleen and benefit Qi, preventing heat-clearing and dampness-draining ingredients from harming the digestive system. Overall, these Southern cough prescriptions are light and flowing in nature, primarily consisting of sweet-cold and bland-leaching herbs while strictly avoiding warm-drying substances. This is consistent with the traditional Southern medical style of using light, clear, and dispersing medications.

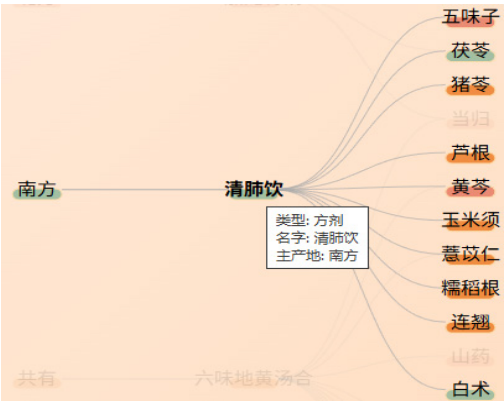


Figure 4. Herbal Composition of *Qingfei Yin*

In contrast, the common Northern prescriptions represented in the red region of the visualization focus on warming Yang to consolidate the foundation and harmonizing the Ying and Wei systems. This approach aligns with the physiological background of the Northern dry-cold climate, where the body is more vulnerable to cold pathogens and damage to Yang Qi. Taking "*Erjia Longgu Tang*" as an example, its formulation clearly reflects the academic emphasis of Northern practitioners on the vulnerability of Yang Qi to cold. The prescription employs pungent and hot ingredients, such as Zingiber officinale (*Shengjiang*), to warm Yang and dispel cold. By pairing Paeonia lactiflora (*Baishao*) with ginger, it establishes a dynamic balance between dispersing and astringing, which harmonizes the Ying and Wei systems, as illustrated in Figure 5. This method skillfully manages complex clinical mechanisms such as "cold-wrapped heat." Overall, Northern prescriptions are characterized by bold, heavy, and potent medicinal properties with an emphasis on warming and tonifying Yang, demonstrating a medication philosophy that stands in sharp contrast to the moistening and clearing methods of the South.



Figure 5. Herbal Composition of *Erjia Longgu Tang*

In summary, the comparative analysis of the formulations of *Qingfei Yin* and *Erjia Longgu Tang* reveals that Southern prescriptions are primarily characterized by "clearing and moistening," whereas Northern prescriptions prioritize "warming and transforming." These differences are not only visually evident in the specific medicinal compositions but also fundamentally permeate the entire process of Pattern Differentiation and Treatment, including the understanding of etiology and pathogenesis, the establishment of therapeutic principles, and the selection of medicinal properties and flavors.

7. Conclusion

This study designed and implemented a visual analytics system centered on the "Disease- Physician-Prescription-Herb" association to explore regional differences in TCM clinical practices. By integrating multi-source medical records and herbal data and leveraging multi-view linkage techniques, the system enables the visual presentation of prescription patterns and herbal composition differences between Northern and Southern China. This system transforms regional clinical experiences into a structured knowledge framework that is summarizable, interactive, and analytical, providing a transparent and reproducible cross-regional tool for TCM clinical decision-making and academic research. Future work will further integrate medical records from diverse TCM schools and incorporate artificial intelligence and knowledge graph technologies. These efforts aim to deepen the intelligent mining and analysis of diagnostic patterns and the complex associations across the "Disease-Syndrome-Treatment-Prescription-Herb" spectrum, ultimately promoting the revitalized heritage, precise dissemination, and innovative development of TCM knowledge in the contemporary era.

Acknowledgments

The authors gratefully acknowledge the financial supported by Special Project for High-quality Development in Key Fields of Zigong First People's Hospital in 2024 (No. 2024 ZDLY02).

References

- [1] XIA Hui-ru. The Thought of "Treatment Based on Syndrome Differentiation" and "Concept of Holism" in Shuowen Jiezi [J]. Lishizhen Medicine and Materia Medica Research, 2019, 30(7): 1696-1697.
- [2] HUANG Zhi-jie. Textual Research on Disease Syndromes and Treatment Characteristics in Northern and Southern Regions [J]. Correspondence Journal of Traditional Chinese Medicine, 2000, (2): 13-14.
- [3] LI Peng, LI Yan, JIN Jia-lin, et al. Visualization Analysis of Hot Trends in Traditional Chinese Medicine Syndrome Research of Coronary Heart Disease Based on CiteSpace [J]. Tianjin Journal of Traditional Chinese Medicine, 2024, 41(8): 1012-1017.
- [4] PENG Xi-chen, LI Ting-bao, YANG Peng-fei. A Systematic Review of Complex Network Associations in Medication Rules of Oral Compound Formulas in Chen Xiuyuan's "Yi Xue Shi Zai Yi" [J]. Journal of Practical Traditional Chinese Internal Medicine, 2018, 32(9): 16-18.
- [5] WANG Song, LI Zheng-jun, YANG Tao, et al. Current Status and Development Trend of Knowledge Graph Research in Traditional Chinese Medicine [J]. Journal of Nanjing

- University of Traditional Chinese Medicine, 2022, 38(3): 272-278.
- [6] REN Lei, DU Yi, MA Shuai, et al. Big Data Visual Analytics Review [J]. Journal of Software, 2014, 25(9): 1909-1936.
 - [7] LIU Qiao, LI Yang, DUAN Hong, et al. Knowledge Graph Construction Techniques [J]. Journal of Computer Research and Development, 2016, 53(3): 582-600.
 - [8] ZHOU Zhi-guang, SHI Chen, SHI Lin-song, et al. A Review of Visual Analysis for Geospatial Data [J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(5): 747-763.
 - [9] GAO Jun. Visualization of Geospatial Data [J]. Engineering of Surveying and Mapping, 2000, (3): 1-7.
 - [10] REN Lei, DU Yi, MA Shuai, et al. A Survey on Big Data Visual Analytics [J]. Journal of Software, 2014, 25(9): 1909-1936.
 - [11] MA Qiu-mei, JIANG Xing-tao, ZENG Qi-feng, et al. Visual Analysis of Sichuan Opera Culture Big Data [J]. Journal of Computer-Aided Design & Computer Graphics, 2025, 37(8): 1465-1476.
 - [12] WANG Fu-qiang. Research on the Theory and Method of Spatial Knowledge Map Construction [D]. PLA Information Engineering University, 2013.
 - [13] HU Hua-hua. Research on the Visual Design of the Style Evolution of Ci Poetry in the Song Dynasty [D]. Harbin Institute of Technology, 2018.
 - [14] ZHANG Jing-zhi, KE Zong-gui. Essence of National Famous TCM Medical Case Records: Diabetes[M]. Guangzhou: Sun Yat-sen University Press, 2019.
 - [15] WANG Yong-yan. Essence of Modern Famous Chinese Medical Cases (Volumes 1-6)[M]. Beijing: People's Medical Publishing House, 2010.
 - [16] HAN Tao. Formulology of Traditional Chinese Medicine[M]. Jinan: Shandong Science and Technology Press, 2020.
 - [17] YANG Huan, YAN Guo-jun, WANG Wei, et al. Chinese Materia Medica[M]. Zhenjiang: Jiangsu University Press, 2023.
 - [18] LIN Yu-lin. Newly Compiled Practical Illustrations of Chinese Medicinal Materials from the Chinese Pharmacopoeia[M]. Beijing: Hualing Press, 2020.