Contents lists available at ScienceDirect

Digital Chinese Medicine

journal homepage: http://www.keaipublishing.com/dcmed

Interdisciplinary integration and development trends of intelligent diagnosis in traditional Chinese medicine: a topic evolution analysis

Chenggong Xie[a], Keying Huang[b], Zhengquan Du[b], Xinyi Huang[b], Bin Wang[a*]

a. Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China

b. Chinese Medicine Science and Technology Cooperation Center, China Academy of Chinese Medical Sciences, Beijing 100700, China

A R T I C L E I N F O

Article history
Received 19 October 2025
Accepted 06 January 2026
Available online 25 March 2026

Keywords
Traditional Chinese medicine diagnosis
Artificial intelligence
Interdisciplinary integration
Research stage identification
Topic evolution analysis
Latent Dirichlet allocation model

A B S T R A C T

Objective To systematically characterize the developmental trajectory and interdisciplinary integration of intelligent diagnosis in traditional Chinese medicine (TCM) through quantitative topic evolution analysis, we addressed the fragmentation of existing research and clarified the long-term research structure and evolutionary patterns of the field.

Methods A topic evolution analysis was performed on Chinese-language literature pertaining to intelligent diagnosis in TCM. Publications were retrieved from the China National Knowledge Infrastructure (CNKI), Wanfang Data, and China Science and Technology Journal Database (VIP), covering the period from database inception to July 3, 2025. A hybrid segmentation approach, based on cumulative publication growth trends and inflection point detection, was applied to divide the research timeline into distinct stages. Subsequently, the latent Dirichlet allocation (LDA) model was used to extract research topics, followed by alignment and evolutionary analysis of topics across different stages.

Results A total of 3 919 publications published between 2003 and 2025 were included, and the research trajectory was divided into five stages based on data-driven breakpoint detection. The field exhibited a clear evolutionary shift from early rule-based systems and tongue-pulse image and signal analysis (2006 – 2010), to machine-learning-based syndrome and prescription modeling (2011 – 2015), followed by deep-learning-driven pattern recognition and formula association (2016 – 2020). Since 2021, research has increasingly emphasized knowledge-graph construction, multimodal integration, and intelligent clinical decision-support systems, with recent studies (2024 – 2025) showing the emergence of large language models and agent-based diagnostic frameworks. Topic evolution analysis further revealed sustained cross-stage continuity in syndrome modeling and prescription association analysis, alongside the progressive consolidation of integrated intelligent diagnostic platforms.

Conclusion By identifying key technological transitions and persistent core research themes, our findings offer a structured reference framework for the design of intelligent diagnostic systems, the construction of knowledge-driven clinical decision-support tools, and the alignment of AI models with TCM diagnostic logic. Importantly, the stage-based evolutionary insights derived from this analysis can inform future methodological choices, improve model interpretability and clinical applicability, and support the translation of intelligent TCM diagnosis from experimental research to real-world clinical practice.

*Corresponding author: Bin Wang, E-mail: wangbin135@139.com.

Peer review under the responsibility of Hunan University of Chinese Medicine.

DOI: 10.1016/j.dcmed.2026.02.001

Citation: XIE CG, HUANG KY, DU ZQ, et al. Interdisciplinary integration and development trends of intelligent diagnosis in traditional Chinese medicine: a topic evolution analysis. Digital Chinese Medicine, 2026, 9(1): 43-56.

Copyright © 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

1 Introduction

Artificial intelligence (AI) technology is being applied increasingly in medicine, and its development continues to advance rapidly [1, 2]. With ongoing AI innovations and broader implementation, the field of traditional Chinese medicine (TCM) has gradually begun to explore intelligent and digital paths [3, 4]. Intelligent diagnosis of TCM is at the intersection of traditional medical practice and modern technology, representing an important breakthrough in the modernisation of TCM. It has gradually become a prominent research highlight in this field [5, 6]. Intelligent TCM diagnosis builds on the four diagnostic methods (inspection, auscultation and olfaction, inquiry, and palpation), and emphasizes treatment based on syndrome differentiation [7]. The process integrates the subjective clinical experience of practitioners with the objective capabilities of intelligent technology developed by researchers.

In recent years, substantial research has accumulated in this area, spanning the construction of syndrome differentiation models based on machine learning (ML), tongue image recognition, pulse signal processing, knowledge graph construction, and other technical directions [8, 9]. With this rapid development, research outputs in TCM intelligent diagnosis have expanded at an unprecedented pace over the past two decades [10]. However, a comprehensive overview of the field's long-term evolutionary trajectory remains lacking.

Existing literature predominantly concentrates on isolated subtopics, such as tongue image analysis, pulse waveform modelling, or syndrome classification, leaving the overall structure and historical evolution of the field unclear [11]. Furthermore, the driving forces behind major technological transitions, including the shift from rule-based systems to ML, and the subsequent progression to deep learning (DL) and multimodal models, have yet to be examined through quantitative analysis. This gap limits the understanding of how technical innovations, interdisciplinary integration, and clinical needs jointly shape the progress of intelligent TCM diagnosis.

In addition, a large body of research on intelligent TCM diagnosis has been published in Chinese-language databases [12], which remains underrepresented in existing quantitative syntheses. Incorporating these studies is essential for capturing the field's authentic developmental trajectory and for accurately identifying its evolutionary patterns and research priorities. Therefore, this study employs an approach integrating bibliometric analysis and topic-modeling-based evolution analysis to systematically organize the research literature in this domain. It aims to delineate the research stages, core themes, and evolutionary pathways, thereby providing a theoretical foundation and decision-making support for future research planning and technological integration.

2 Data and methods

2.1 Search strategy

This study retrieved publications indexed in the China National Knowledge Infrastructure (CNKI), Wanfang Data, and China Science and Technology Journal Database (VIP). To comprehensively capture literature on intelligent diagnosis in TCM, the final retrieval was constructed by combining two themes: "TCM diagnosis" and "AI". The literature search encompassed records from database inception to July 3, 2025. The complete search strategy is detailed in Supplementary Table S1.

2.2 Study selection

Inclusion criteria: (i) the study explicitly focused on TCM diagnostic processes, including but not limited to syndrome differentiation, diagnostic reasoning based on the four diagnostic methods, tongue diagnosis, pulse diagnosis, constitution identification, or integrated disease-syndrome diagnosis; (ii) AI technologies were directly applied to support, simulate, or optimize TCM diagnostic tasks, rather than solely addressing treatment recommendation, drug discovery, or pharmacological analysis; (iii) the publication is an original research article.

Exclusion criteria: (i) publications not written in Chinese; (ii) non-research articles (e.g., reviews, editorials, conference abstracts, and news reports); (iii) studies lacking essential bibliographic information, such as author list, publication year, or abstract; (iv) studies in which AI techniques were applied only to TCM treatment, prescription optimization, drug compatibility analysis, or pharmaceutical research, without explicit involvement of diagnostic reasoning or diagnostic decision-making.

2.3 Data extraction and preprocessing

Two independent reviewers with a background in TCM diagnostics and AI, performed screening according to predetermined criteria. To ensure data quality and consistency, a multi-stage preprocessing pipeline was applied. First, essential fields, including title, first author, publication year, and abstract, were extracted from the original Excel files. Records containing missing values in these fields were excluded. Next, Chinese word segmentation was performed using a regular expression pattern to extract tokens consisting of two or more consecutive Chinese characters. To enhance the relevance of extracted terms, a customized TCM-specific stop-word lexicon was applied. This lexicon included non-informative words, common methodological terms, high-frequency but semantically generic TCM terms, and TCM terms not directly related to diagnostic content. This step aimed to reduce noise and improve the discriminative capacity of the retained terms.

2.4 Research stage identification

To examine the relationship between the developmental process of intelligent diagnosis in TCM and its temporal evolution, we sought to partition the overall time span into distinct chronological phases. While related studies have commonly employed single-year intervals as temporal units [13], the selection of an appropriate segmentation scale requires further justification when the observed evolution spans a short or extended period [14, 15]. In this study, we adopted a hybrid segmentation strategy combining inflection-point detection and fixed-interval anchoring.

We constructed a time-series dataset comprising annual publication counts and their cumulative totals. A segmentation pipeline integrating four complementary methods was employed to identify potential research stages. First, second-order differentiation of the cumulative publication curve was used to detect shifts in acceleration [16]. Second, k-means clustering was applied normalized temporal features to identify structural transitions [17]. Third, the detection of abnormal year-over-year growth-rate changes were detected and curvature analysis was performed based on smoothed trends derived from moving averages. Detected inflection points were then merged, deduplicated, and filtered. To enhance interpretability, fixed five-year interval boundaries (2010, 2015, and 2020) were incorporated as additional reference points. Finally, adjacent points within two years were merged to avoid over-segmentation, and the total number of stage boundaries was limited to four.

2.5 Research topic extraction and evolution analysis

We conducted a topic evolution analysis to uncover underlying thematic patterns in the literature and trace the emergence, evolution, and decline of topics over time. First, we performed topic extraction. Basic methods such as term frequency (TF) and TF-inverse document frequency (IDF) help identify important and high-frequency keywords, providing an initial overview of key research terms [18]. However, these methods lack semantic depth and thematic coherence. To address this limitation, we employed topic modeling, an unsupervised ML technique widely used for discovering abstract topics within large text corpora [19]. Among available models, we selected the latent Dirichlet allocation (LDA) model for topic extraction [20].

Then, we divided the corpus according to the previously defined research stages, training separate LDA models on each stage. Finally, topics were aligned across stages using similarity metrics (e.g., cosine similarity), enabling us to map topic trajectories and identify emerging, stable, and declining themes [21]. A complete flowchart illustrating each part of the study is presented in Figure 1.

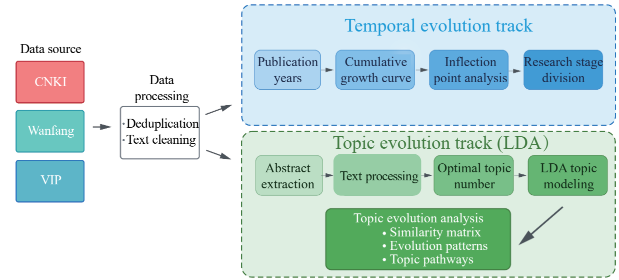

Figure 1 Framework of literature analysis of TCM intelligent diagnosis

All of the above analyses were conducted using Python 3.13. Specifically, both Pandas and NumPy were employed for data handling and numerical operations, as well as Matplotlib for visualising trends and inflection points, along with Regex for processing text patterns [22]. Topic modelling and clustering were performed using scikit-learn (Count Vectorizer, LDA, and k-means) [23].

3 Results

3.1 Literature retrieval and screening

A total of 15 824 records were initially retrieved, including 3 697 from CNKI, 7 605 from Wanfang, and 4 522 from VIP. Following screening, 1 265, 2 327, and 855 publications were included from CNKI, Wanfang, and VIP, respectively. After removing duplicated and unrelated records, 3 919 studies were retained for subsequent data analysis.

3.2 Research stage

Due to the extremely low volume of publications before 2003, we designated 2003 as the starting point for stage segmentation, with the endpoint set at the current year (2025). As described earlier, we applied a mixed-methods segmentation pipeline to identify distinct research stages. Specifically, second-order difference detection identified one breakpoint, clustering identified two, growth-rate analysis identified one, and moving average analysis revealed nine.

After removing duplicates, ten candidate breakpoints were obtained, including 2004, 2013, 2017, 2018, 2019, 2020, 2021, 2022, 2023, and 2024. Among these, 2022, 2023, and 2024 were identified as the most prominent and were selected as data-driven segmentation markers. In addition, three fixed time points—2010, 2015, and 2020—were incorporated based on established knowledge of the field's developmental trajectory. Consequently, the final segmentation points were determined to be 2010, 2015, 2020, and 2023.

For completeness, the period before 2006 (2003 – 2005) was considered as a separate early stage. Although this period comprised a non-negligible number of publications (n = 118, 3.0%), the related studies were predominantly exploratory in nature, focusing on conceptual discussions, narrative reviews, and preliminary methodological proposals rather than digital modelling or algorithmic applications. Accordingly, topic evolution analysis was conducted for the period (2006 – 2025). The final stage divisions are presented in Table 1, and the segmentation outcomes together with the final stage boundaries are illustrated in Figure 2.

Table 1 Temporal distribution of publications across research stages in TCM intelligent diagnosis from 2003 to 2025

Research stage (year)	Number of publications	Proportion (%)
Pre-2005 (2003 – 2005)	118	3.0
Stage 1 (2006 – 2010)	407	10.5
Stage 2 (2011 – 2015)	445	11.5
Stage 3 (2016 – 2020)	1 076	27.7
Stage 4 (2021 – 2023)	1 313	33.8
Stage 5 (2024 – 2025)	522	13.5

Data for 2025 included publications indexed up to July 3, 2025.

To clearly illustrate the TCM intelligent diagnostic system, we established a system framework comprising four layers: a data layer, a knowledge layer, an algorithm layer, and an application layer (Figure 3).

3.3 Research topics extraction

We applied the LDA model to extract research topics from the abstracts of papers in each stage. Abstracts were chosen as the corpus because they provide a concise yet informative summary of each study while retaining essential contextual and semantic details [24]. Their structured format also facilitates preprocessing tasks, such as tokenization and stop-word removal [25].

Following text preprocessing, we extracted eight topics per research stage. For each topic, the top eight representative keywords were identified. A complete summary of all topics and their associated keywords is provided in Supplementary Table S2, with a set of eight representative topics across five research stages.

In stage 1, topic 1 focused on clinical application, with emphasis on medication use, symptom patterns, and

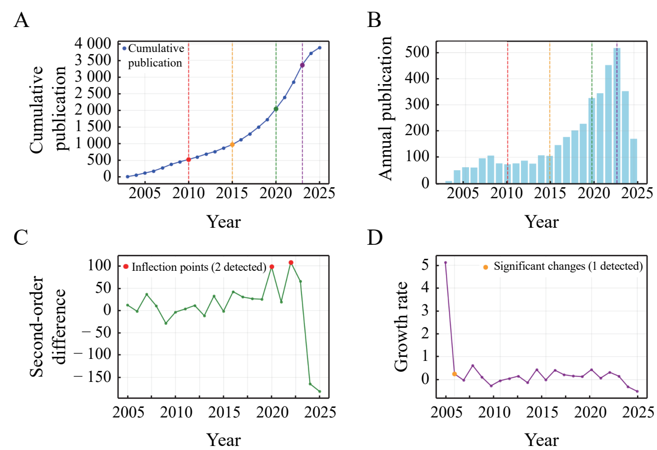

Figure 2 Growth analysis of TCM intelligent diagnosis research

A, cumulative publication growth curve with segmentation points. B, annual publication distribution across defined research stages. C, inflection point detection using second-order differential analysis. D, publication growth rate and identification of significant change points. Data for 2025 included publications indexed up to July 3, 2025.

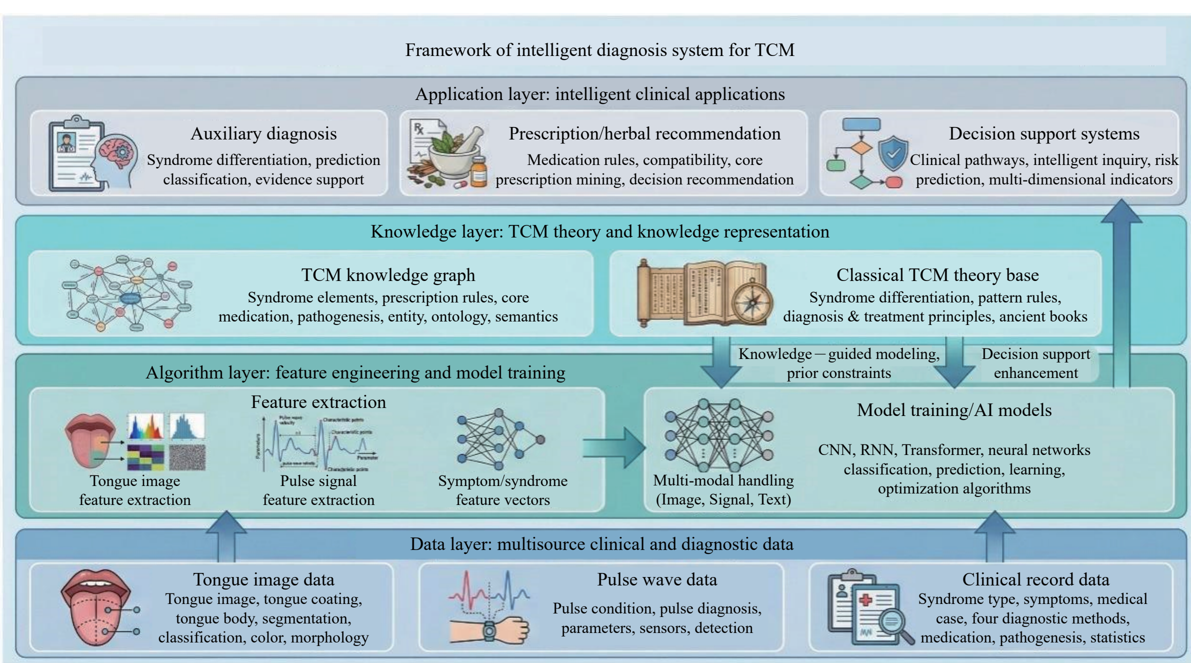

Figure 3 Framework of TCM intelligent diagnosis system

therapeutic principles. Topic 2 focused on pattern differentiation and knowledge discovery through data mining and algorithmic approaches. Topic 3 addressed the classification and standardization of tongue images, highlighting visual features and objective quantification. Topic 4 explored the integration of neural networks with traditional diagnostic frameworks, with an emphasis on algorithmic learning from clinical cases. Topic 5 captured syndrome differentiation and standardization efforts, including the identification of diagnostic elements and expert consensus. Topic 6 concentrated on image processing techniques for tongue diagnosis, particularly segmentation, fuzzy systems, and regional analysis. Topic 7 investigated the quantification of pulse characteristics through sensor-based systems and signal-processing methods. Topic 8 elaborated on changes in tongue features under pathological conditions, focusing on statistical analysis and clinical grouping.

In stage 2, topic 1 highlighted modeling approaches for treatment evaluation, with emphasis on efficacy assessment, control-group comparison, and statistical significance. Topic 2 focused on pathomechanism analysis and symptom association through cases data. Topic 3 centered on prescriptions, featuring core drug combinations and compatibility strategies. Topic 4 reflected diagnostic modeling of pulse conditions, integrating algorithms, classification techniques, and knowledge structures. Topic 5 captured neural network-based syndrome prediction, structured data utilization, and diagnostic system integration. Topic 6 emphasized tongue diagnosis image processing, including segmentation algorithms and feature extraction of tongue characteristics. Topic 7 explored clustering and quantification of symptoms and syndromes, with a focus on syndrome elements and frequency patterns. Topic 8 dealt with expert-driven syndrome differentiation, combining structured data, symptom profiles, and integrated diagnostic schemes.

In stage 3, topic 1 involved symptom association analysis in case records, focusing on patterns, syndromes, and medication rules. Topic 2 explored syndrome differentiation modeling using predictive algorithms and neural networks. Topic 3 centered on pulse diagnosis, emphasizing signal parameters, feature recognition, and objective assessment. Topic 4 highlighted the integration of diagnostic knowledge and data-driven systems to support clinical decision-making. Topic 5 reflected advances in tongue image analysis, including segmentation, classification, and ML techniques. Topic 6 emphasized prescription pattern discovery, capturing herbal combinations, core prescriptions, and knowledge inheritance. Topic 7 focused on syndrome element quantification and expert-based feature extraction. Topic 8 investigated statistical modeling of syndrome factors, including intervention comparisons, correlation analysis, and efficacy evaluation.

In stage 4, topic 1 focused on medication association analysis, highlighting prescriptions, core combinations, and symptom-based rules derived from clinical records. Topic 2 centered on expert-driven diagnostic and treatment systems, emphasizing the integration of domain knowledge, clinical guidelines, and AI. Topic 3 involved statistical analysis of syndrome types and elements, incorporating demographic and comorbidity factors. Topic 4 emphasized intervention studies, including control-group comparisons, scoring systems, and outcome assessments. Topic 5 explored knowledge graph (KG) construction, linking classical texts, clinical data, and decision-support recommendations. Topic 6 highlighted molecular mechanisms and network pharmacology, with a focus on targets, pathways, and active ingredients. Topic 7 investigated syndrome differentiation modeling using machine learning, predictive algorithms, and variable analysis. Topic 8 centered on tongue image analysis, covering classification, segmentation, and algorithm development.

In stage 5, topic 1 centered on syndrome prediction modeling, statistical analysis integration, tongue diagnosis, and influencing factors. Topic 2 focused on medication association analysis, emphasizing syndrome types, herbal combinations, and rule-based knowledge mining. Topic 3 highlighted tongue image classification, involving ML, image segmentation, and diagnostic modeling. Topic 4 addressed KG construction, covering symptom extraction, clinical cases, ontologies, and semantic representation. Topic 5 related to bibliometric analysis, examining publication trends, institutional collaboration, and research highlights. Topic 6 emphasized intelligent diagnosis and treatment systems, featuring AI, four diagnostic methods, and clinical decision-making. Topic 7 explored multidimensional system measurement, including gene markers, acupuncture point selection, and evaluation models. Topic 8 pertained to clinical trials, involving control-group comparisons, intervention outcomes, statistical analysis, and syndrome differentiation.

3.4 Research topic evolution

To clearly illustrate the dominant AI technologies, applications, and thematic evolutionary characteristics across each stage, we compiled the summary presented in Table 2.

As described above, certain topics persist across stages, while others emerge or fade over time. To characterize patterns of topic evolution, we calculated cosine similarity among topics in adjacent stages. A similarity score above 0.5 indicates meaningful thematic continuity, while scores below this threshold imply topic divergence or disappearance. This approach enables us to trace the emergence, persistence, and decline of topics across the research timeline.

Table 2 Evolution of AI technologies in TCM intelligent diagnosis from 2006 to 2025

Stage	Year	Dominant AI technology	Primary application in TCM diagnosis	Representative keyword	Cross-technology evolution feature	Evolution trend
Stage 1	2006 to 2010	Rule based systems, early ML including support vector machines, decision trees, shallow neural networks, tongue image segmentation, pulse signal parameter extraction	Tongue image objectification, pulse waveform detection, rule based syndrome differentiation, case based reasoning	Rule, data mining, neural network, segmentation, pulse condition, tongue image	Transition from experience based reasoning to data driven models, initial progress in objective tongue and pulse analysis	Transition from handcrafted rules toward data oriented modeling as digital TCM records began to accumulate
Stage 2	2011 to 2015	Enhanced ML including support vector machines, random forests, k-nearest neighbours, statistical modelling, medical case knowledge extraction, advanced image processing	Detailed tongue and coating analysis, symptom to pattern association, knowledge extraction and formula association, pulse signal analysis	Data mining, classification, tongue diagnosis, pulse condition, medical case	Wider adoption of ML. More mature image algorithms, statistical models increasingly used in pattern differentiation	Expansion of digitized TCM data enables shift from shallow models to deep representation learning
Stage 3	2016 to 2020	DL with convolutional and recurrent models, feature learning, large scale image driven models, early development of knowledge structured resources	DL based tongue classification, pulse pattern recognition, syndrome prediction, formula association analysis	Classification, learning, tongue image, correlation, prediction, pulse condition	Marked shift from ML to DL, tongue image modelling becomes central, structured knowledge resources develop more rapidly	Rapid growth of multimodal TCM datasets accelerates integration of visual and physiological signals
Stage 4	2021 to 2023	Knowledge graphs, multimodal learning, hybrid models that integrate knowledge and data, intelligent reasoning systems	Multimodal pattern identification, knowledge guided diagnostic support, treatment recommendation, combined tongue and pulse models	Knowledge, entity, diagnosis and treatment, model, domain, AI	Convergence of DL and knowledge based reasoning, growth of multimodal fusion and knowledge enhanced interpretability	Increasing demand for unified diagnostic frameworks drives fusion of symbolic knowledge and learned representations
Stage 5	2024 to 2025	Large language models, multimodal foundation models, agent based intelligent systems, self supervised representation learning	Complex reasoning in pattern identification, intelligent formula generation with clinical explanations, unified representation of tongue pulse and symptoms, automated diagnostic workflows	Intelligent system, map, multidimensional measurement, learning, syndrome differentiation	Transformation toward agent driven intelligence, strongest level of technique convergence, unified multimodal representations and advanced inference capabilities become prominent	Emergence of TCM oriented foundation models indicates movement toward general purpose intelligent diagnostic systems

To visualize thematic shifts between research stages, we present four complementary figures. The topic similarity heatmap illustrates the semantic alignment between topics of preceding and subsequent stages. The similarity distribution plot illustrates overall semantic coherence, distinguishing unrelated pairs with a threshold. The maximum similarity plot indicates which topics from an earlier stage continue into the following stage. Finally, the topic evolution paths diagram illustrates how topics diverge, converge, or remain stable across stages. Figure 4 – 7 collectively summarize topic evolution between adjacent research phases.

For conciseness, stages and topics are abbreviated as S*n*T*n* in the following sections (e.g., stage 1, topic 2 is denoted S1T2). In Figure 4, strong thematic associations are observed between S1T6 and S2T6, with similarity values reaching or exceeding 0.8. Similarity scores are largely distributed between 0.1 and 0.3, and the maximum similarity for each topic exceeds 0.3.

In Figure 5, strong associations are identified

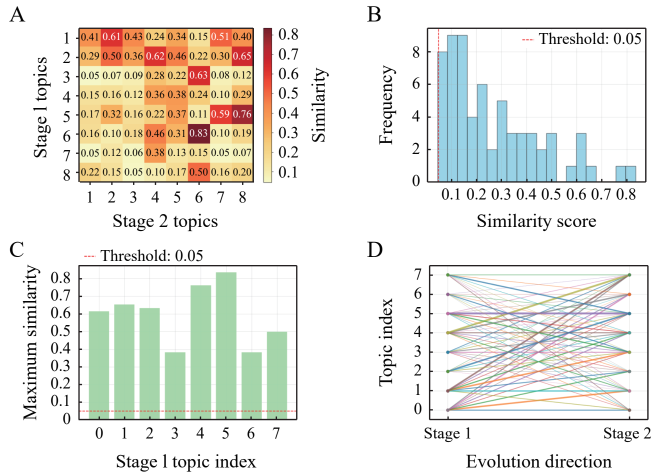

Figure 4 Research on the topic evolution analysis from stage 1 to stage 2

A, cosine similarity heatmap between topics of adjacent stages. B, distribution of all pairwise similarity scores. C, maximum similarity for each topic from the preceding stage. D, directed graph of topic evolution paths.

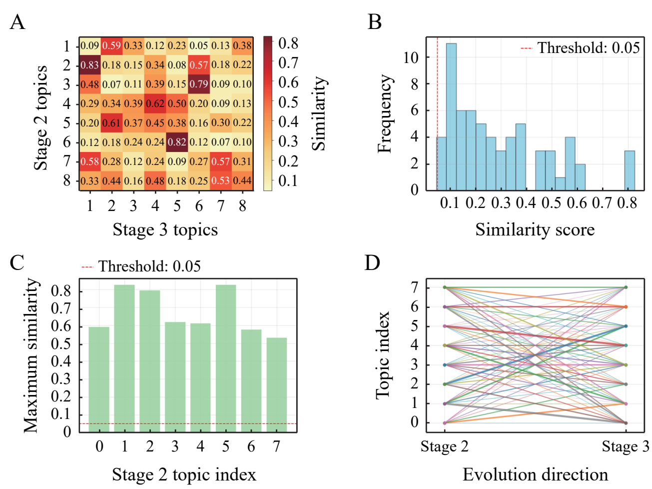

Figure 5 Research on the topic evolution analysis from stage 2 to stage 3

A, cosine similarity heatmap between topics of adjacent stages. B, distribution of all pairwise similarity scores. C, maximum similarity for each topic from the preceding stage. D, directed graph of topic evolution paths.

between S2T2 and S3T1, as well as between S2T6 and S3T5, with similarity values of at least 0.8. The similarity scores are mainly concentrated between 0.1 and 0.4, and the maximum similarity for each topic is greater than 0.5.

In Figure 6, strong associations are evident between S3T1 and S4T1, S3T6 and S4T1, S3T8 and S4T3, S3T2 and S4T7, and S3T5 and S4T8, each with similarity values of at least 0.8. Similarity scores are predominantly distributed between 0.1 and 0.4, and the maximum similarity for every topic exceeds 0.3.

In Figure 7, strong associations are observed between S4T1 and S5T2, S4T8 and S5T3, and S4T5 and S5T4, with similarity values equal to or exceeding 0.8. Similarity scores are largely concentrated between 0.1 and 0.4, and the maximum similarity for each topic exceeds 0.2.

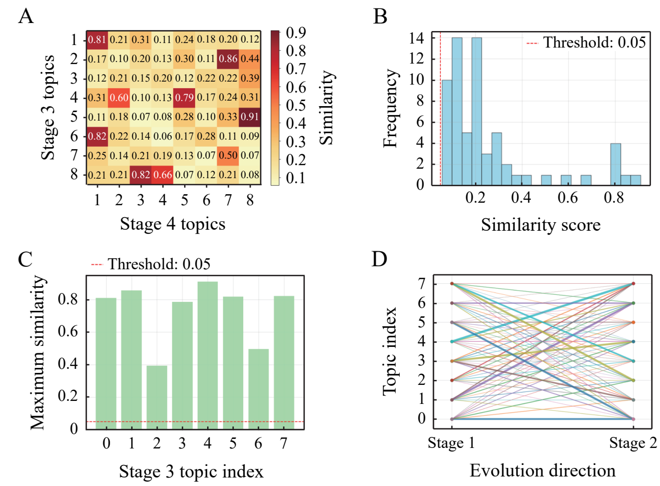

Figure 6 Research on the topic evolution analysis from stage 3 to stage 4

A, cosine similarity heatmap between topics of adjacent stages. B, distribution of all pairwise similarity scores. C, maximum similarity for each topic from the preceding stage. D, directed graph of topic evolution paths.

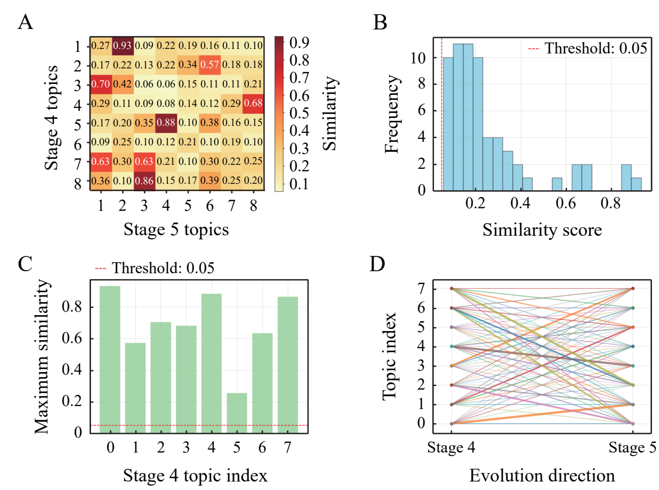

Figure 7 Research on the topic evolution analysis from stage 4 to stage 5

A, cosine similarity heatmap between topics of adjacent stages. B, distribution of all pairwise similarity scores. C, maximum similarity for each topic from the preceding stage. D, directed graph of topic evolution paths.

Overall, similarity values between topics in adjacent stages are mainly distributed between 0.1 and 0.4, indicating moderate to relatively weak associations, while all topics remain above the predefined relevance threshold. A limited number of topic pairs exhibit similarity values above 0.8, reflecting strong thematic continuity.

4 Discussion

4.1 Research findings and thematic evolution analysis

This study provides the comprehensive, quantitative mapping of the developmental stages, thematic structures, and technological evolution in the field of intelligent diagnosis for TCM. By integrating bibliometric analysis with longitudinal topic modeling, we delineated a research trajectory spanning 2006 to 2025 and identify five distinct stages that reflect shifts in methodologies, data modalities, and interdisciplinary integration. The timeline is divided as follows.

(i) Emerging phase (pre-2005): characterized by sparse publications and nascent research activity.

(ii) Early exploration stage 1 (2006 – 2010): initial integration of data mining and neural networks based on traditional diagnostic elements, alongside feasibility exploration of interdisciplinary approaches.

(iii) Technical expansion stage 2 (2011 – 2015): maturation of AI techniques such as data mining and modelling, which facilitated advances in standardisation and objectivity, along with early efforts toward multidimensional data integration.

(iv) Expansion and integration stage 3 (2016 – 2020): marked by substantial data growth and increasing algorithmic complexity, with a focus on multi-source data correlation and practical diagnostic models.

(v) Intelligent refinement stage 4 (2021 – 2023): deep integration of AI and biotechnology, driven by KG and multi-omics approaches, promoting interdisciplinary and fine-grained research.

(vi) Frontier science stage 5 (2024 – 2025): research shifts toward practical intelligent systems, comprehensive data integration, and multidisciplinary collaboration, extending beyond the AI-TCM dual framework.

In stage 1 (2006 – 2010), research focused on the objectification of traditional diagnostic indicators such as tongue and pulse features, including tongue image segmentation and pulse-signal parameter extraction [26, 27]. Early applications of data mining and neural networks were explored to support syndrome classification and medication pattern analysis [28, 29].

From stage 1 to stage 2, the transition involved 641 shared keywords and 8 topic links, with cosine similarities ranging from 0.046 to 0.835. The strongest continuity was observed from S1T6 to S2T6 (similarity 0.835), both centered on tongue diagnosis, algorithms, and imaging [30]. Early studies on tongue segmentation and coating recognition (S1T6) laid the groundwork for automated tongue-analysis systems (S2T6) [31, 32].

The most pronounced cross topic evolution was observed from S1T5 to S2T8 (similarity 0.761), reflecting a shift from syndrome standardization to expert-system research [33, 34]. Three topic splits and one merge were identified. S1T1 diverged into S2T2 and S2T7, with drug-symptom association studies evolving into case-based drug mining and symptom quantifications. S1T2 split into S2T4 and S2T8, representing an expansion from syndrome mining toward pulse diagnosis and expert-knowledge analysis [35]. S1T5 diverged into S2T7 and S2T8, reflecting the development of syndrome standardization into syndrome quantification and expert-system research [36]. S1T3 and S1T6 merged into S2T6, integrating tongue-imaging techniques with diagnostic algorithms [37].

In stage 2 (2011 – 2015), improvements in tongue- and pulse-recognition algorithms were documented [38]. Medical case mining was introduced to identify medication-compatibility patterns and to establish quantitative links among syndrome differentiation, symptoms, and therapeutic outcomes [39].

From stage 2 to stage 3, 704 shared keywords and 10 topic links were identified, with similarity scores ranging from 0.046 to 0.825. This period reflected a shift from technical integration toward large-scale diagnostic applications [40]. The two strongest evolutionary paths, S2T2 to S3T1 and S2T6 to S3T5, demonstrated clear thematic continuity and stable progression. S2T2, which focused on mining drug-symptom relationships, evolved into S3T1's broader analysis of clinical correlations [41, 42]. Similarly, S2T6, originally centered on tongue-image segmentation, developed into S3T5's DL methods for diagnostic tasks [43].

S3T1 integrated insights from S2T2 and S2T7, combining clinical rule mining with symptom quantification [44]. S3T2 synthesized modeling approaches from S2T1 and S2T5, emphasizing predictive tools for syndrome differentiation [45]. Meanwhile, S2T2 diverged into both S3T1 and S3T6, with the latter focusing on drug-knowledge transfer [46]. S2T3 transformed into S3T6, and S2T6 continued into S3T5, signaling a transition from algorithm development to practical diagnostic applications [47]. S2T7 further branched into S3T1 and S3T7, indicating diversification in the study of symptom and syndrome elements [48].

In stage 3 (2016 – 2020), intelligent recognition systems for tongue and pulse diagnosis matured considerably, leading to the development of standardized diagnostic tools [49]. Large-scale medical record analyses were conducted to reveal intrinsic associations between symptom clusters and syndromes, enabling quantification of syndrome components [50, 51].

From stage 3 to stage 4, 806 shared terms and 8 topic links were identified, with similarity scores ranging from 0.056 to 0.909. Research shifted from large-scale diagnostic modeling to deeper integration of AI and multi-omics approaches [52]. The strongest continuity was observed from S3T5 to S4T8 (similarity 0.909), reflecting the evolution from tongue image analysis to intelligent diagnostic systems. S3T2 to S4T7 (similarity 0.856) also showed strong alignment, continuing the work on syndrome prediction [53].

S4T1 drew comparably from S3T1 (similarity 0.810) and S3T6 (similarity 0.817), integrating clinical-rule mining and drug knowledge transfer. Both S4T2 and S4T5 stemmed from S3T4 (similarity 0.602 and 0.786), advancing structured diagnostic knowledge, including KG frameworks [54]. Topic divergence was evident in the split of S3T8 into S4T3 (similarity 0.821) and S4T4 (similarity 0.664), forming distinct research lines focusing on population-level syndrome features and outcome evaluation, respectively [55].

In stage 4 (2021 – 2023), KGs were employed to integrate classical literature and clinical data, facilitating the construction of structured TCM knowledge systems [56]. ML models were widely adopted for syndrome prediction, medication recommendation, and mechanistic studies linking herbal medicine to molecular targets [57].

From stage 4 to stage 5, 748 shared keywords and 8 topic links were identified, with similarity scores ranging from 0.057 to 0.932. The transition involved 837 shared terms and 8 topic links, with similarity scores ranging from 0.054 to 0.932. The strongest link was observed from S4T1 to S5T2 (similarity 0.932), reflecting a continued focus on drug-rule patterns and intelligent prescription analysis [58].

Topic merging occurred as S5T3 drew from S4T8 and S4T7, combining tongue-diagnosis techniques with model optimization approaches [59]. S5T1 integrated insights from S4T3 and S4T7, linking syndrome modeling and population-based analysis. Meanwhile, S4T7 diverged into S5T1 and S5T3, showing divergence into two distinct directions: syndrome modeling and image-based learning [60, 61].

In stage 5 (2024 – 2025), intelligent diagnostic systems expanded to encompass all four diagnostic methods [5]. Multidimensional data were incorporated into syndrome analysis [62]. Emphasis was placed on the development of standardized databases and cross-institutional collaborative networks.

4.2 Research thinking

4.2.1 Integration of TCM diagnostic philosophy and AI paradigms The core of TCM diagnosis rests on the holistic concept, syndrome differentiation, and individualized treatment. The holistic concept emphasizes systemic relationships, dynamic balance, and functional interactions rather than isolated pathological changes. Syndrome differentiation focuses on complex constellations of symptoms and etiological heterogeneity rather than single diagnostic labels. Individualized treatment highlights patient-centered decision-making rather than outcomes driven solely by data or AI models.

Contemporary AI paradigms have begun to approximate these principles to varying degrees. Multimodal learning partially reflects the holistic concept by integrating information from the four diagnostic methods [63]. Syndrome-label prediction corresponds to a formalized expression of syndrome differentiation [64]. Personalized modeling and few-shot learning approaches attempt to address individual variability in clinical presentation. However, these efforts remain fragmented and localized. No existing AI framework fully embodies the integrative logic of TCM diagnostic thinking.

Most current studies reduce TCM diagnosis to a classification task, prioritizing data standardization and objectification while weakening theoretical depth. The subjectivity inherent in TCM diagnosis is not a methodological flaw but reflects accumulated clinical experience and contextual judgment. In contrast, AI-driven objectification relies on predefined labels, rules, and data consistency. Achieving a balance between experiential subjectivity and algorithmic objectivity remains a central theoretical and methodological challenge for TCM intelligent diagnosis research.

4.2.2 Progress and methodological characteristics of AI models The evolution of intelligent diagnosis in TCM reflects the co-development of AI technologies and the deepening integration between computational paradigms and TCM diagnostic principles. This progression represents a systematic adaptation to changing data availability, methodological capacity, clinical needs, and interdisciplinary contexts. Each stage emerged in response to the structural limitations of the preceding phase.

During the early exploration stage, the absence of large-scale digitized clinical data and standardized diagnostic criteria led to a strong reliance on expert-defined rules and mappings between syndromes, treatments, and outcomes. Consequently, rule-based systems and shallow models became feasible tools for diagnostic reasoning [65]. Early attempts at intelligent tongue and pulse diagnosis were exploratory and limited in scale [66]. These approaches demonstrated restricted capacity to address diagnostic complexity, scalability, and interpretability, thereby driving a natural transition toward data-driven methods.

During the technical expansion stage, the widespread adoption of electronic medical records and the accumulation of structured TCM data enabled the shift from rule-based reasoning to statistical learning. ML models offered greater adaptability and the ability to uncover empirical patterns beyond explicit expert knowledge [67]. In tongue and pulse diagnosis, emphasis shifted toward feature engineering and quantitative representation [68, 69]. Classical ML methods also provided partial interpretability through explanation techniques. However, their dependence on manually designed features constrained their ability to model high-dimensional diagnostic data—a limitation that directly motivated the adoption of DL approaches.

During the expansion and integration stage, DL models enabled automatic feature learning and substantially reduced reliance on manual annotation [70]. These models were effectively applied to large-scale tongue images, tongue coating data, and pulse waveforms, supporting end-to-end diagnostic modeling. Although diagnostic performance improved markedly, limited interpretability and weak alignment with TCM theory reduced clinical trust. This disconnect prompted efforts to integrate data-driven learning with structured TCM knowledge, shaping a distinctive developmental pathway for intelligent TCM diagnosis.

During the intelligent refinement stage, the research focus shifted from performance optimization to interpretability, consistency, and clinical usability. KGs and multi-modal models were incorporated to embed TCM theory, standardized diagnostic criteria, and expert experience into AI systems [71, 72]. Tongue and pulse diagnosis increasingly relied on integrated reasoning that synergized data with structured knowledge. Despite enhanced clinical applicability, these approaches remained predominantly task-specific and exhibited limited generalizability across diverse TCM diagnostic contexts, underscoring the need for more universal intelligent paradigms.

In the frontier science stage, the emergence of LLMs has driven a paradigm shift from task-oriented modeling toward general-purpose intelligent systems. These models support representation learning, complex reasoning, and adaptive interaction, which conceptually aligns with the holistic concept in TCM [58]. Research emphasis has shifted toward comprehensive systems capable of knowledge reasoning, clinical explanation, and decision support across diverse diagnostic tasks [73, 74]. Non-reasoning LLMs exhibit limited interpretability, which constrains their clinical application. Reasoning models and hybrid models can present explicit inference processes and enhance transparency; however, their internal mechanisms remain partially opaque and prone to hallucination risks [75].

4.2.3 Challenges in clinical translation and real-world implementation Although research on intelligent diagnosis in TCM has advanced rapidly, a substantial gap persists between methodological development and clinical translation in real-world settings. The deployment of intelligent TCM diagnostic systems is not merely an AI technical challenge, but a complex problem involving systematization and institutionalization. This gap is primarily driven by inadequate clinical validation, limitations in data sharing and standardization, and constraints at the institutional and ecosystem levels.

Most existing studies rely on retrospective or single-center datasets, lacking prospective, multicenter, and real-world datasets. Consequently, the clinical applicability of many proposed models remains limited. In addition, foundational research in AI models has largely prioritized objective performance metrics such as accuracy, precision, sensitivity, specificity, the area under the receiver operating characteristic (ROC) curve (AUC), and F1-score, while issues pertaining to clinical usability and decision reliability have received insufficient attention.

Data sharing mechanisms remain underdeveloped due to concerns over patient privacy and institutional barriers, leading to duplicated efforts and reduced research efficiency. Furthermore, multiple versions of TCM diagnostic standards coexist in clinical practice [76]. The acquisition and processing of diagnostic data, particularly tongue and pulse information, lack unified technical specifications. Heterogeneity in data standards across studies impedes large-scale integration and comparative analysis.

Additional challenges stem from the disciplinary divide between clinicians and algorithm developers, exacerbated by a shortage of interdisciplinary professionals with expertise in both domains. The rapid pace of AI innovation has outpaced the development of corresponding regulatory frameworks, leaving issues related to medical supervision, accountability, and ethical governance insufficiently defined [77]. Collaboration among academia, industry, and clinical institutions remains in its infancy, limiting the effective translation of research outcomes. Collectively, these challenges constrain both the advancement and clinical adoption of intelligent TCM diagnostic systems.

4.3 Research limitations

This study has several limitations. First, the identification of literature related to TCM intelligent diagnosis relied on manual screening, which may have resulted in data being omitted or misclassified. Second, the analysis focused on quantity rather than quality, failing to account for factors such as journal impact, academic influence, or study design. Third, due to the complexity of TCM diagnostic terms and the associated challenges in translation and data processing, English-language publications were excluded. Consequently, some relevant studies may not have been fully represented.

4.4 Research directions and future priorities

Future research on TCM intelligent diagnosis should transition from task-specific modeling toward integrated and generalizable systems that are more deeply aligned with TCM theory and validated using large-scale, multicenter, and real-world data.

First, future models should support cross-modal reasoning by integrating tongue diagnosis, pulse diagnosis, symptoms, medical history, constitution, and patient demographic data. Such integration is essential for generating coherent diagnostic interpretations instead of isolated predictions.

Second, priority should be given to the deep integration of TCM theory and explainable AI. Embedding TCM diagnostic reasoning, syndrome structure, and expert decision-making pathways into model construction is critical for clinical applicability. Explainability should be consistent with real-world TCM diagnostic workflows, rather than relying on generic post hoc explanations.

Third, data standardization and sharing should be prioritized. Unified standards for data acquisition, annotation, and preprocessing, together with mechanisms for open data sharing across institutions and research groups, are necessary to enable large-scale integration, comparative analysis, and the reduction of redundant efforts.

Fourth, effective translation of research outcomes requires systematic governance of TCM intelligent diagnostic systems. This includes the refinement of regulatory frameworks, policy support, and the strengthening of collaboration among academia, industry, and clinical institutions etc.

Finally, sustained innovation in the field depends on enhanced interdisciplinary collaboration among clinicians, data scientists, engineers, and policymakers. Fostering interdisciplinary talent with dual expertise in TCM and AI is critical to ensuring responsible innovation and the effective clinical translation of TCM intelligent diagnostic technologies.

5 Conclusion

Our study analyzed the developmental trajectory, research hotspots, and thematic evolution in the field of TCM intelligent diagnosis. Over the past two decades, the field has progressed from foundational research to the development of integrated diagnostic systems. The early phases emphasized the objectification of traditional diagnostic indicators and explored the feasibility of AI methods. Subsequently, the introduction of large-scale clinical data, advanced modeling technologies, and standardized diagnostic frameworks characterized the middle stage. More recently, research has increasingly integrated KGs, multi-omics, and DL, enabling intelligent systems to support more refined and structurally grounded diagnostic processes. As the field continues to evolve, future efforts are likely to focus on the development of clinically validated tools, enhanced interpretability of AI models, and deeper integration of TCM theory with biomedical evidence.

Fundings

Grants of National Natural Science Foundation of China (82274685).

Author contributions

Chenggong Xie: conceptualization, investigation, data curation, methodology, formal analysis, and writing – original draft. Keying Huang and Zhengquan Du: investigation and data curation. Xinyi Huang: formal analysis and validation. Bin Wang: methodology, supervision, writing – review & editing, and project administration. All authors approved the submission and take responsibility for this manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

[1] DEO RC. Artificial intelligence and machine learning in cardiology. Circulation, 2024, 149(16): 1235–1237.

[2] KHOSRAVI P, FUCHS TJ, HO DJ. Artificial intelligence-driven cancer diagnostics: enhancing radiology and pathology through reproducibility, explainability, and multimodality. Cancer Research, 2025, 85(13): 2356–2367.

[3] LIN YM, ZHANG Y, WANG DY, et al. Computer especially AI-assisted drug virtual screening and design in traditional Chinese medicine. Phytomedicine, 2022, 107: 154481.

[4] ZHANG YL, SUN WH, YANG CG, et al. TCMP-300: a comprehensive traditional Chinese medicinal plant dataset for plant recognition. Scientific Data, 2025, 12: 1166.

[5] TIAN DC, CHEN WH, XU DC, et al. A review of traditional Chinese medicine diagnosis using machine learning: inspection, auscultation-olfaction, inquiry, and palpation. Computers in Biology and Medicine, 2024, 170: 108074.

[6] WANG L, TANG KQ, WANG Y, et al. Advancements in artificial intelligence-driven diagnostic models for traditional Chinese medicine. The American Journal of Chinese Medicine, 2025, 53(3): 647–673.

[7] TIAN ZK, WANG DJ, SUN X, et al. Current status and trends of artificial intelligence research on the four traditional Chinese medicine diagnostic methods: a scientometric study. Annals of Translational Medicine, 2023, 11(3): 145.

[8] ZHUANG Y, YU LK, JIANG N, et al. TCM-KLLaMA: intelligent generation model for traditional Chinese medicine prescriptions based on knowledge graph and large language model. Computers in Biology and Medicine, 2025, 189: 109887.

[9] YUAN L, YANG L, ZHANG S, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. EClinicalMedicine, 2023, 57: 101834.

[10] SONG ZL, CHEN GX, CHEN CY. AI empowering traditional Chinese medicine? Chemical Science, 2024, 15(41): 16844–16886.

[11] LI QL, LIU Z. Tongue color analysis and discrimination based on hyperspectral images. Computerized Medical Imaging and Graphics, 2009, 33(3): 217–221.

[12] SUN WY, WANG R, OUYANG S, et al. “Weibing” in traditional Chinese medicine: biological basis and mathematical representation of disease-susceptible state. Acta Pharmaceutica Sinica B, 2025, 15(5): 2363–2371.

[13] HE XF, PENG C, XU YX, et al. Global scientific research landscape on medical informatics from 2011 to 2020: bibliometric analysis. JMIR Medical Informatics, 2022, 10(4): e33842.

[14] XIN ML, BI FJ, WANG C, et al. The circadian rhythm: a new

target of natural products that can protect against diseases of the metabolic system, cardiovascular system, and nervous system. Journal of Advanced Research, 2025, 69: 495–514.

[15] SALTINI M, VASCONCELOS P, RUEFFLER C. Complex life cycles drive community assembly through immigration and adaptive diversification. Ecology Letters, 2023, 26(7): 1084–1094.

[16] YANG MH, XU D, CUI QW, et al. An efficient fisher matrix approximation method for large-scale neural network optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(5): 5391-5403.

[17] PEI SF, CHEN HM, NIE FP, et al. Centerless clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 167–181.

[18] OZYEGEN O, KABE D, CEVIK M. Word-level text highlighting of medical texts for telehealth services. Artificial Intelligence in Medicine, 2022, 127: 102284.

[19] ENNAJARI H, BOUGUILA N, BENTAHAR J. Correlated topic modeling for short texts in spherical embedding spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(6): 4567–4578.

[20] WILCOX KT, JACOBUCCI R, ZHANG ZY, et al. Supervised latent Dirichlet allocation with covariates: a Bayesian structural and measurement model of text and covariates. Psychological Methods, 2023, 28(5): 1178–1206.

[21] PAL R, CHOPRA H, AWASTHI R, et al. Predicting emerging themes in rapidly expanding COVID-19 literature with unsupervised word embeddings and machine learning: evidence-based study. Journal of Medical Internet Research, 2022, 24(11): e34067.

[22] HARRIS CR, MILLMAN KJ, VAN DER WALT SJ, et al. Array programming with NumPy. Nature, 2020, 585(7825): 357–362.

[23] LINDELÖF G, ALEDAVOOD T, KELLER B. Dynamics of the negative discourse toward COVID-19 vaccines: topic modeling study and an annotated data set of twitter posts. Journal of Medical Internet Research, 2023, 25: e41319.

[24] KHAN MS, SHAIKH A, OCHANI RK, et al. Assessing the quality of abstracts in randomized controlled trials published in high impact cardiovascular journals. Circulation: Cardiovascular Quality and Outcomes, 2019, 12(5): e005260.

[25] GUO E, GUPTA M, DENG JW, et al. Automated paper screening for clinical reviews using large language models: data analysis study. Journal of Medical Internet Research, 2024, 26: e48996.

[26] ZHI L, ZHANG D, YAN JQ, et al. Classification of hyperspectral medical tongue images for tongue diagnosis. Computerized Medical Imaging and Graphics, 2007, 31(8): 672–678.

[27] WILLUM HANSEN T, STAESSEN JA, TORP-PEDERSEN C, et al. Prognostic value of aortic pulse wave velocity as index of arterial stiffness in the general population. Circulation, 2006, 113(5): 664–670.

[28] ZHOU XZ, CHEN SB, LIU BY, et al. Development of traditional Chinese medicine clinical data warehouse for medical knowledge discovery and decision support. Artificial Intelligence in Medicine, 2010, 48(2/3): 139–152.

[29] CHEN YJ, LIU YY, ZHAO GZ, et al. Chinese traditional medicine recognition by support vector machine (SVM) terahertz spectrum. Spectroscopy and Spectral Analysis, 2009, 29(9): 2346–2350.

[30] LI QL, WANG YT, LIU HY, et al. Sublingual vein extraction algorithm based on hyperspectral tongue imaging technology. Computerized Medical Imaging and Graphics, 2011, 35(3): 179–185.

[31] WANG XZ, ZHANG B, YANG ZM, et al. Statistical analysis of tongue images for feature extraction and diagnostics. IEEE Transactions on Image Processing, 2013, 22(12): 5336–5347.

[32] YAN ZF, WANG KQ, LI NM. Computerized feature quantification of sublingual veins from color sublingual images. Computer Methods and Programs in Biomedicine, 2009, 93(2): 192–205.

[33] ZHANG NL, YUAN SH, CHEN T, et al. Latent tree models and diagnosis in traditional Chinese medicine. Artificial Intelligence in Medicine, 2008, 42(3): 229–245.

[34] LAM CFD, LEUNG KS, HENG PA, et al. Chinese acupuncture expert system (CAES): a useful tool to practice and learn medical acupuncture. Journal of Medical Systems, 2012, 36(3): 1883–1890.

[35] CHUNG CY, CHENG YW, LUO CH. Neural network study for standardizing pulse-taking depth by the width of artery. Computers in Biology and Medicine, 2015, 57: 26–31.

[36] LIN YY, XUE YY, YU J, et al. A quantification model of traditional Chinese medicine syndromes in children with idiopathic precocious puberty and early puberty. Journal of Traditional Chinese Medicine, 2013, 33(5): 630–636.

[37] CHUNG YF, HU CS, YEH CC, et al. How to standardize the pulse-taking method of traditional Chinese medicine pulse diagnosis. Computers in Biology and Medicine, 2013, 43(4): 342–349.

[38] WANG HY, WANG X, DELLER JR, et al. Shape-preserving preprocessing for human pulse signals based on adaptive parameter determination. IEEE Transactions on Biomedical Circuits and Systems, 2014, 8(4): 594–604.

[39] LI JL. Combination of symptoms, syndrome and disease: treatment of refractory diabetic gastroparesis. World Journal of Gastroenterology, 2014, 20(26): 8674.

[40] LIU BY. Utilizing big data to build personalized technology and system of diagnosis and treatment in traditional Chinese medicine. Frontiers of Medicine, 2014, 8(3): 272–278.

[41] CHEN HY, LIN YH, WU JC, et al. Prescription patterns of Chinese herbal products for menopausal syndrome: analysis of a nationwide prescription database. Journal of Ethnopharmacology, 2011, 137(3): 1261–1266.

[42] LIU XJ, LV M, WANG YZ, et al. Deciphering the compatibility rules of traditional Chinese medicine prescriptions based on NMR metabolomics: a case study of Xiaoyaosan. Journal of Ethnopharmacology, 2020, 254: 112726.

[43] LI XQ, ZHANG Y, CUI Q, et al. Tooth-marked tongue recognition using multiple instance learning and CNN features. IEEE Transactions on Cybernetics, 2019, 49(2): 380–387.

[44] HE RP, JIN Z, MA RY, et al. Network pharmacology unveils spleen-fortifying effect of Codonopsis Radix on different gastric diseases based on theory of “same treatment for different diseases” in traditional Chinese medicine. Chinese Herbal Medicines, 2021, 13(2): 189–201.

[45] LIU JY, LI YY, WEI LL, et al. Screening and identification of potential biomarkers and establishment of the diagnostic serum proteomic model for the traditional Chinese medicine syndromes of tuberculosis. Journal of Ethnopharmacology, 2014,

155(2): 1322–1331.
[46] DUAN DD, WANG Z, ZHANG BL, et al. Fangjiomics: revealing adaptive omics pharmacological mechanisms of the myriad combination therapies to achieve personalized medicine. Acta Pharmacologica Sinica, 2015, 36(6): 651–653.
[47] WEN GH, MA JJ, HU Y, et al. Grouping attributes zero-shot learning for tongue constitution recognition. Artificial Intelligence in Medicine, 2020, 109: 101951.
[48] WANG JK, MA QT, LI YQ, et al. Research progress on traditional Chinese medicine syndromes of diabetes mellitus. Biomedicine & Pharmacotherapy, 2020, 121: 109565.
[49] HAO YM, CHENG F, PHAM M, et al. A noninvasive, economical, and instant-result method to diagnose and monitor type 2 diabetes using pulse wave: case-control study. JMIR mHealth and uHealth, 2019, 7(4): e11959.
[50] HU QN, YU T, LI JH, et al. End-to-end syndrome differentiation of Yin deficiency and Yang deficiency in traditional Chinese medicine. Computer Methods and Programs in Biomedicine, 2019, 174: 9–15.
[51] LIU LJ, LIU L, FU XD, et al. A cloud-based framework for large-scale traditional Chinese medical record retrieval. Journal of Biomedical Informatics, 2018, 77: 21–33.
[52] LI X, REN J, ZHANG W, et al. LTM-TCM: a comprehensive database for the linking of traditional Chinese medicine with modern medicine at molecular and phenotypic levels. Pharmacological Research, 2022, 178: 106185.
[53] ZHANG H, NI WD, LI J, et al. Artificial intelligence-based traditional Chinese medicine assistive diagnostic system: validation study. JMIR Medical Informatics, 2020, 8(6): e17608.
[54] YU T, LI JH, YU Q, et al. Knowledge graph for TCM health preservation: design, construction, and applications. Artificial Intelligence in Medicine, 2017, 77: 48–52.
[55] FU C, ZHANG NL, CHEN BX, et al. Identification and classification of traditional Chinese medicine syndrome types among senior patients with vascular mild cognitive impairment using latent tree analysis. Journal of Integrative Medicine, 2017, 15(3): 186–200.
[56] LIU C, LI Z, LI JM, et al. Research on traditional Chinese medicine: domain knowledge graph completion and quality evaluation. JMIR Medical Informatics, 2024, 12: e55090.
[57] LV QJ, CHEN GX, HE HH, et al. TCMBank-the largest TCM database provides deep learning-based Chinese-Western medicine exclusion prediction. Signal Transduction and Targeted Therapy, 2023, 8: 127.
[58] YANG K, DONG X, ZHANG SH, et al. PresRecRF: herbal prescription recommendation via the representation fusion of large TCM semantics and molecular knowledge. Phytomedicine, 2024, 135: 156116.
[59] HU Y, WEN GH, LIAO HQ, et al. Automatic construction of Chinese herbal prescriptions from tongue images using CNNs and auxiliary latent therapy topics. IEEE Transactions on Cybernetics, 2021, 51(2): 708–721.
[60] ABLAT N, ABLIMIT M, SUN Y, et al. Application of new imaging methods in the development of Chinese medicine. Biomedicine & Pharmacotherapy, 2022, 153: 113470.
[61] YUE WJ, JI WD, WANG XY, et al. SDPR: prescription recommendation with syndrome differentiation in traditional Chinese medicine. IEEE Journal of Biomedical and Health Informatics, 2025, 29(5): 3736–3749.
[62] HUANG L, WANG Q, DUAN QC, et al. TCMSSD: a comprehensive database focused on syndrome standardization. Phytomedicine, 2024, 128: 155486.
[63] WANG J, LIU YM, LI J, et al. Artificial intelligence in traditional Chinese medicine: multimodal fusion and machine learning for EnhancedDiagnosis and TreatmentEfficacy. Current Medical Science, 2025, 45(5): 1013–1022.
[64] LIM J, LI JY, FENG X, et al. Predicting TCM patterns in PCOS patients: an exploration of feature selection methods and multi-label machine learning models. Heliyon, 2024, 10(15): e35283.
[65] ZHANG H, ZHU X. TCM diagnosis and reasoning method based on grey relational analysis. Journal of Biomedical Engineering, 2007, 24(1): 206–209.
[66] ANASTASI JK, CURRIE LM, KIM GH. Understanding diagnostic reasoning in TCM practice: tongue diagnosis. Alternative Therapies in Health and Medicine, 2009, 15(3): 18–28.
[67] ZHAO CB, LI GZ, WANG CJ, et al. Advances in patient classification for traditional Chinese medicine: a machine learning perspective. Evidence-Based Complementary and Alternative Medicine, 2015, 2015: 376716.
[68] ZUO WM, WANG P, ZHANG D. Comparison of three different types of wrist pulse signals by their physical meanings and diagnosis performance. IEEE Journal of Biomedical and Health Informatics, 2016, 20(1): 119–127.
[69] HU MC, CHENG MH, LAN KC. Color correction parameter estimation on the smartphone and its application to automatic tongue diagnosis. Journal of Medical Systems, 2015, 40(1): 18.
[70] LIANG ZH, LIU J, OU AH, et al. Deep generative learning for automated EHR diagnosis of traditional Chinese medicine. Computer Methods and Programs in Biomedicine, 2019, 174: 17–23.
[71] ZHAO XL, WANG YF, LI PH, et al. The construction of a TCM knowledge graph and application of potential knowledge discovery in diabetic kidney disease by integrating diagnosis and treatment guidelines and real-world clinical data. Frontiers in Pharmacology, 2023, 14: 1147677.
[72] WENG H, CHEN JL, OU AH, et al. Leveraging representation learning for the construction and application of a knowledge graph for traditional Chinese medicine: framework development study. JMIR Medical Informatics, 2022, 10(9): e38414.
[73] HUA R, DONG X, WEI Y, et al. Lingdan: enhancing encoding of traditional Chinese medicine knowledge for clinical reasoning tasks with large language models. Journal of the American Medical Informatics Association, 2024, 31(9): 2019–2029.
[74] LI YX, ELNAFFAR S, CHEN HY, et al. An LLM method for understanding traditional Chinese medicine: mechanism exploration and innovative application. IEEE Journal of Biomedical and Health Informatics, 2025: 1–14.
[75] JIANG QY, HUANG WH, LIANG JF, et al. A novel intelligent model for visualized inference of medical diagnosis: a case of TCM. Artificial Intelligence in Medicine, 2024, 149: 102799.
[76] WU P, LI J, YAN HX, et al. Status and prospect of international standardization of TCM diagnosis. Pharmacological Research, 2021, 171: 105746.
[77] MENG QQ, MI Y, WANG F, et al. Intelligent technology leads the transformation of traditional Chinese medicine: large models and virtual cells aid modern analysis of stroke treatment. Pharmacological Research, 2025, 221: 107953.

(Editor-in-Charge Jie Deng)

中医智能诊断的跨学科融合与发展趋势：主题演变分析

谢成功[a], 黄可盈[b], 杜政泉[b], 黄馨懿[b], 王斌[a*]

a. 中国中医科学院中医药信息研究所, 北京 100700, 中国
b. 中国中医科学院中医药科技合作中心, 北京 100700, 中国

【**摘要**】**目的** 本研究旨在通过定量主题演变分析，系统阐述中医智能诊断研究的发展脉络及跨学科融合特征，以整合现有分散的研究，揭示该领域的长期研究结构与演化规律。**方法** 本研究对中医智能诊断的中文文献进行了主题演变分析。从中国知网、万方和维普数据库中检索相关文献，时间范围从数据库创建到 2025 年 7 月 3 日。基于文献累积发文增长趋势与拐点检测相结合的混合分期方法，对研究时间线进行阶段划分。随后利用隐含狄利克雷分配（LDA）模型提取研究主题，并对不同阶段的研究主题进行比对与演化分析。**结果** 本研究共纳入 2003—2025 年发表的相关文献 3 919 篇，并根据数据驱动的断点检测方法，将研究轨迹划分为 5 个阶段。研究主题整体呈现出清晰的演进路径：从早期的基于规则的系统和舌脉图像和信号分析（2006—2010 年），到基于机器学习的证候与方药建模（2011—2015 年），然后是深度学习驱动的模式识别和方剂关联（2016—2020 年）。自 2021 年以来，研究重点逐渐转向知识图谱构建、多模态信息融合及智能临床决策支持系统，且在 2024—2025 年间，出现了以大语言模型和智能体为代表的新型诊断框架。主题演变分析进一步揭示了证候建模和处方关联分析的持续跨阶段连续性，以及集成智能诊断平台的逐步整合。**结论** 本研究通过识别关键技术转折点与长期稳定的核心研究主题，为智能中医诊断系统的设计、知识驱动型临床决策支持工具的构建，以及人工智能模型与中医诊断思维的对齐提供了结构化参考框架。基于阶段划分的演化认知有助于指导未来方法学选择，提升模型的可解释性和临床适用性，并推动中医智能诊断由实验研究向真实临床应用转化。

【**关键词**】中医诊断；人工智能；跨学科融合；研究阶段识别；主题演变分析；隐含狄利克雷分配模型