Contents lists available at ScienceDirect

Digital Chinese Medicine

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

Construction of the clinical diagnosis and treatment model during the "pre-disease to disease" window period in traditional Chinese medicine: integration of objective multimodal data from the perspective of traditional Chinese medicine stateology

Danyang Li[a], Min Ai[b], Pai Zhou[b], Ying Deng[b], Chaoyang Yang[a], Qinghua Peng[b, c*]

a. School of Traditional Chinese Medicine, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian 350122, China

b. School of Traditional Chinese Medicine, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China

c. Department of Ophthalmology, The First Hospital of Hunan University of Chinese Medicine, Changsha, Hunan 410007, China

A R T I C L E I N F O

Article history

Received 15 March 2026

Accepted 06 April 2026

Available online 25 June 2026

Keywords

Traditional Chinese medicine stateology

Window period

Clinical diagnostic and therapeutic model

Multimodal data fusion

Deep learning

A B S T R A C T

The philosophy of "treating disease before its onset" is a fundamental concept of traditional Chinese medicine (TCM), permeating its diagnostic and therapeutic framework, and is central to clinical practice. However, current TCM diagnostic and treatment models for the "pre-disease to disease" window period face several limitations, including the lack of comprehensive clinical parameters, difficulties in characterizing and integrating heterogeneous multimodal data, and insufficient dynamic precision in interventions and efficacy evaluations. To address these issues, guided by Professor Candong Li's theory of TCM stateology, this study focuses on integrating objective multimodal data. It proposes a new model for personalized TCM diagnosis and treatment targeting the "pre-disease to disease" window period. This approach first proposes the idea of restructuring the conceptual framework of "symptom" and integrating multi-source heterogeneous data at macroscopic, mesoscopic, and microscopic levels to form a three-dimensional assessment indicator system. By integrating graph neural networks, convolutional neural networks, attention mechanisms, and knowledge graph-guided weight allocation, this approach enables collaborative representation, alignment, and fusion of multi-source data. Subsequently, it plans to construct a multimodal fusion model at both feature and decision levels, in order to establish mappings between indicators and TCM state elements, and to screen key indicators characterizing pathological evolution during the window period. Furthermore, it proposes a technical path for enhancing model interpretability using methods such as SHapley Additive exPlanations (SHAP) and Ablation-CAM++. Finally, with state assessment as the core, it proposes the concept of constructing a dynamic evaluation method for individualized diagnosis and treatment based on time-series data analysis using algorithms such as long short-term memory (LSTM) networks and gated recurrent units (GRUs). Moreover, a causal inference framework and semi-supervised learning strategies are introduced to enable quantitative evaluation of individual intervention effects and to provide interpretable therapeutic feedback, forming a complete technical path from data representation and fusion, weight adjustment, and interpretability analysis, to dynamic diagnosis feedback. This study aims to address deficiencies in the current TCM diagnosis and treatment model during the "pre-disease to disease" window period and to provide an operational framework for the clinical practice of TCM's "treating disease before its onset".

*Corresponding author: Qinghua Peng, E-mail: pqh410007@126.com.

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

DOI: 10.1016/j.dcmed.2026.05.002

Citation: LI DY, AI M, ZHOU P, et al. Construction of the clinical diagnosis and treatment model during the "pre-disease to disease" window period in traditional Chinese medicine: integration of objective multimodal data from the perspective of traditional Chinese medicine stateology. Digital Chinese Medicine, 2026, 9(2): 173-183.

1 Introduction

In the context of China's Healthy China Strategy, which emphasizes shifting the focus of disease prevention forward, the clinical value of the traditional Chinese medicine (TCM) concept of "treating disease before its onset" has become more apparent [1]. The "pre-disease to disease" window period, as an important stage in disease occurrence and progression, represents both a turning point in pathological evolution and an important window for early intervention to achieve effective disease prevention. At present, research on identifying subhealthy states and early biomarkers of diseases has advanced considerably. Furthermore, the rapid development of objective data acquisition technologies based on the TCM four diagnostic methods has facilitated the extensive collection of multimodal data covering the entire disease process [2, 3]. However, limited by the disease-centered paradigm of modern medicine, current studies on the "pre-disease to disease" window period remain insufficiently precise. Current research predominantly focuses on single-dimensional data and lacks integrated analyses of multidimensional and multimodal data. Consequently, they cannot adequately capture the complex and dynamic pathological features and evolutionary patterns of the window period, hindering the development of individualized TCM clinical intervention paradigms based on dynamic evolutionary patterns.

The TCM stateology theory proposed by Professor Candong Li focuses on states of human life activities throughout the life cycle, encompassing health, pre-disease, disease, and post-disease stages. By analyzing temporal and spatial changes during state transitions, this theory provides a methodological framework for investigating pathological evolution during the window period [4]. Guided by TCM stateology, this study introduces technologies for multimodal data representation and integration. By synthesizing heterogeneous multi-source data at macroscopic, mesoscopic, and microscopic levels, this research constructs an individualized TCM clinical diagnosis and treatment model that focuses on state identification during the window period. This approach aims to provide a reference paradigm grounded in original TCM thought for clinical practice in "treating disease before its onset".

2 Basic connotation of the "pre-disease to disease" window period and the guiding significance of TCM stateology

The concept of the window period initially appeared in the World Health Organization (WHO) guidelines for infectious disease management, referring to the time from pathogen infection to the detection of infection biomarkers, representing the incubation or subclinical period of disease [5]. From the perspective of TCM, the "pre-disease to disease" window period corresponds to the "pre-disease" stage within the "treating disease before its onset" theory. At this stage, classical TCM manifestations of diseases or syndromes have not yet fully developed; however, concealed pathological changes, such as disharmony of Qi and blood, imbalance of Yin and Yang, and impairment of organ function, are already present [6]. The unique pathological nature of the window period dictates that its diagnosis and treatment cannot be merely a continuation of the static "diagnosis and treatment based on syndrome differentiation" model. Instead, it is essential to establish a new framework capable of identifying and describing "subclinical syndromes" or "pre-syndromic states" to achieve precise, dynamic assessment. The theory of TCM stateology offers methodological guidance in this regard.

"States" constitute the starting point for human health cognition, summarizing physiological functions and pathological changes throughout all disease stages, including the pre-disease stage. This concept encompasses not only syndrome stages but also pre-syndrome stages [7]. Professor Wenfeng Zhu's "syndrome element differentiation" theory divides syndromes into minimal units of disease location and disease nature, quantifying these elements to facilitate objective TCM diagnosis in overt disease stages [8]. Building upon syndrome element differentiation theory, Professor Li integrated stateology to propose "state elements", encompassing dimensions of location, nature, and severity. These elements provide holistic and dynamic identification of principal contradictions across disease phases, from pre-disease to post-disease, enabling precise TCM intervention based on identified location, nature, and severity [9]. Consequently, TCM clinical diagnosis and treatment during the "pre-disease to disease" window period can be effectively evaluated using TCM states. In summary, based on TCM stateology, the dynamic identification of TCM state factors may support precise intervention and assessment during the window period, translating the concept of "preventing disease before its onset" into a quantifiable clinical pathway.

3 Challenges of TCM clinical diagnosis and treatment during the "pre-disease to disease" window period

3.1 Deficiency in evaluation indicators and lack of a multi-dimensional data collection framework

Macroscopic parameters, including geographic, meteorological, and seasonal natural characteristics, are often neglected in assessment systems of the "pre-disease to disease" window period [10]. TCM adheres to the principle that "humans resonate with heaven and earth and correspond to the sun and moon", as described in the *Huangdi*

Neijing (《黄帝内经》, *Inner Canon of Huangdi*): "life arises from the Qi of heaven and earth, shaped by the laws of the four seasons". Human development depends upon the interplay and harmony of Yin-Yang energies influenced by seasonal, geographic, and climatic factors, reflected in diverse lifestyle, dietary, exercise, and psychological states. Therefore, pathological evolution within the body is also a product of specific spatiotemporal conditions, and identifying critical nodes during dynamic pathological changes in the window period is essential for TCM clinical practice. Ignoring macroscopic parameters weakens the spatiotemporal interpretability of diagnostic indicators, transforming individuals into isolated organic entities at disconnected temporal and spatial points, reducing the diagnostic and early-warning efficacy of indicators in dynamic contexts, and compromising the principle of genuine "resonance between humans and nature".

At the mesoscopic level, current research in TCM clinical diagnosis and treatment primarily focuses on collecting general clinical symptoms and diagnostic indicators. However, during the transition from "pre-disease" to "disease", clinical symptoms are often minimal, atypical, or even absent, exhibiting a characteristic of "no evident symptoms to observe" [11]. Historically, this might be attributed to technical limitations that prevented the detection of subtle mesoscopic changes. Nevertheless, with advancements in modern detection technologies, such as pulse diagnosis instruments, tongue imaging systems, electronic noses, and standardized TCM inquiry models, objective collection of pulse waveforms, tongue images, odor information, and inquiry data has become feasible, objectively supporting assessing pathological states in the window period [12-14].

At the microscopic level, such indicators, due to their inherent sensitivity, typically exhibit high diagnostic value during pathological evolution within the window period. However, current research generally relies on single or limited laboratory parameters, capturing only localized pathological changes and failing to reflect the systemic, dynamic pathological imbalance in the body [15]. In addition, the TCM holistic concept emphasizes that "internal pathologies must manifest externally", implying an intrinsic relationship between internal pathological changes and external macroscopic manifestations. Analyzing isolated laboratory indicators detached from their holistic context risks losing connection to the overall TCM state, introducing assessment discrepancies.

3.2 Difficulty in multimodal data fusion and insufficient model interpretability

Although certain macroscopic, mesoscopic, and microscopic indicators have demonstrated potential diagnostic value for window period interventions, current TCM clinical studies often rely on analyses limited to single-modality data [16, 17]. Due to varying standards for objective data and a lack of effective information fusion techniques, information barriers persist across different medical and health data modalities (e.g., textual and image data). Consequently, existing studies rarely provide unified representations of multimodal indicators, limiting the holistic perspective central to TCM and hindering a comprehensive understanding of pathological evolution.

Furthermore, effective fusion of multi-source information is vital for data processing, improving system performance and robustness. Traditional fusion algorithms, such as support vector machines, random forests, and decision tree models, impose high demands on data quality, often struggling to uncover deep-level correlations within high-dimensional, information-rich multimodal TCM datasets. In addition, the reliance on expert-based feature classification and extraction further increases the complexity of their implementation [18]. By contrast, deep learning, characterized by powerful hierarchical, nonlinear fitting capabilities, addresses the bottlenecks encountered by traditional machine learning methods in representing high-dimensional, multimodal data and effectively identifies deep correlations between indicators and TCM states. Nevertheless, the inherent complexity and abstraction in data processing make deep learning models opaque and hard to interpret [19]. Moreover, foundational issues, such as limited TCM corpus resources and insufficient standardization of state identification systems during the window period, limit these models' ability to capture the essence of TCM state differentiation. Consequently, these models risk becoming static labeling tools, hindering interpretability and meaningful exploration of authentic biological implications, restricting their further application in window period clinical diagnosis and treatment.

Lastly, existing research on TCM diagnosis and treatment during the window period predominantly employs cross-sectional methodologies, rarely incorporating time-series data collection. The absence of sequential multimodal data collection hinders precise characterization of pathological evolution during the window period, weakening the close relationships between multimodal indicators under consistent spatiotemporal conditions. Establishing deeper causal connections between indicators and TCM states becomes difficult, introducing spatiotemporal biases into diagnosis, which affects the reliability and stability of diagnostic outcomes [20]. Therefore, when constructing deep learning models for medical diagnosis, careful consideration of spatiotemporal data characteristics is essential for enhancing diagnostic accuracy and interpretability.

3.3 Insufficient precision in intervention efficacy and imperfect real-time evaluation mechanism

Clinically, the window period is the pre-clinical stage of the disease. However, due to its characteristic of having inconspicuous clinical symptoms and atypical syndromes, the current clinical intervention decisions are still mostly based on the model of the disease stage that has already occurred. That is, through syndrome differentiation and treatment, a fixed intervention plan is adopted for a specific syndrome type to achieve personalized, individualized health management strategies. Since this strategy targets the relative stability of established disease stages, frequent adjustments to intervention decisions are typically unnecessary. However, during the dynamic phase of the window period, this fixed intervention plan has a lag effect. It is difficult to provide real-time feedback on the continuous, dynamic changes in the window period. This may result in missed critical opportunities for TCM intervention in the field of TCM diagnosis and treatment [21], leading to the continuous progression of the pathological state into the disease stage [22]. Furthermore, existing intervention decisions are mostly based on periodic assessments of four diagnostic parameters, rather than on real-time monitoring and integration of macroscopic, mesoscopic, and microscopic indicators, limiting the establishment of real-time response mechanisms to pathological changes.

Regarding efficacy evaluation, current assessment systems primarily rely on endpoint outcomes. Although these indicators are essential for evaluating treatment effects at the disease stage, the focus during the window period, as a transitional phase between health and disease, should be on whether pathological indicators improve or whether risk factors for disease progression are reduced, rather than on final disease outcomes. Consequently, relying solely on key outcomes for efficacy assessment risks overlooking process-related changes during the window period, making it difficult to promptly identify and quantify the intervention's true benefits. In addition, the therapeutic effects during the window period cannot be adequately evaluated with a single indicator; instead, they require an integrated assessment of macroscopic, mesoscopic, and microscopic parameters. The existing study has introduced deep learning models and other complex algorithms to evaluate treatment efficacy in the disease stage [23]. However, due to the lack of algorithmic models capable of processing temporal multimodal data and supporting the assessment of dynamic efficacy trajectories, it is difficult to map the trajectory of efficacy evolution accurately. This makes it challenging to objectively capture the true effects of interventions during the window of opportunity and hinders the optimization of personalized treatment strategies based on dynamic feedback.

4 Development strategies for a TCM diagnosis and treatment model in the window period from the perspective of stateology

4.1 Acquisition and representation of “macro-meso-micro” three-dimensional data

4.1.1 Construction of a “macro-meso-micro” three-dimensional acquisition framework Symptoms and clinical manifestations constitute the primary sources of diagnostic information in clinical practice. To address the challenges posed by the “absence of observable symptoms” and the limitations of single-modality information during the window period, this requires expanding the scope of symptom data collection in clinical research. This can be achieved by integrating macroscopic, mesoscopic, and microscopic parameters to establish a three-dimensional spatiotemporal data-acquisition framework for “macro-meso-micro” indicators.

At the macroscopic level, emphasis should be placed on natural environmental and temporal rhythm parameters, including geographical location, seasonal variation, meteorological conditions, and the Wuyun (五运, five movements) and Liuqi (六气, six climates) theory.

At the mesoscopic level, in addition to general clinical information and structured or scale-based four-diagnostic data, modern detection technologies, including fundus cameras, pulse diagnosis instruments, and tongue diagnostic systems, should be employed to collect latent symptoms, referring to mesoscopic indicators that are difficult to capture objectively through conventional diagnostic techniques.

At the microscopic level, in addition to routine biochemical indicators, biological samples such as serum, urine, and feces should be collected. Modern analytical techniques, including proteomics, metabolomics, and gut microbiome analysis, can be applied to identify key molecular expressions associated with pathological changes. Importantly, multimodal data across these three dimensions should cover the entire disease trajectory, from pre-disease to disease and post-disease stages, within a spatiotemporal framework, enabling the characterization of dynamic expression patterns of key indicators during disease development.

4.1.2 Multimodal data representation under the three-dimensional framework Although multimodal data enable a fine-grained characterization of TCM state changes from multiple dimensions, their multi-source and heterogeneous nature significantly increases analytical complexity. Therefore, data preprocessing based on multimodal representation techniques is required for textual, image-based, time-series, and spatio-temporal data across different modalities to extract feature vectors for each modality, to enable a comprehensive and multidimensional representation of the critical transitional state during the window period.

For medical and health data such as general clinical records and structured four diagnostic datasets, techniques including lexical analysis, syntactic parsing, and named entity recognition of medical terminology, combined with entity alignment [24], representation learning [25], and graph neural networks (GNNs) [26], can be applied to explore semantic differences among medical terms and compute vector similarity. This process enables semantic alignment and representation of medical terminology, addressing common issues in medical datasets, such as subjectivity, polysemy, semantic ambiguity, and non-standardized expression.

For visual modalities, including facial images, tongue images, and fundus images, convolutional neural networks (CNN) [27], attention mechanisms (AM) [28], and transfer learning techniques [29] can be integrated to extract features from multiple perspectives, including color and texture, generating reliable feature representations of facial, lingual, and ocular images.

For multi-source time series data, a differential processing strategy is used to preserve, to the greatest extent possible, data from key points in the progression of TCM conditions. For high-frequency, regularly sampled data such as pulse waveforms and blood glucose levels, the original sampling frequency is retained. Time-domain sequence data, along with Fourier- and wavelet-transformed sequences, are used to extract time-domain, frequency-domain, and time-frequency fusion features, forming a sequence of feature vectors. For low-frequency, irregularly sampled data such as biochemical indicators and omics data, the original timestamps are retained, and no forced interpolation is performed; instead, an event-based alignment method (using clinical milestones, such as changes in TCM conditions and syndromes, as reference points) is used to process the non-equidistant sequences directly. Finally, non-linear interpolation fuses multimodal data on a unified timeline, feeding the combined data into a deep learning model to achieve high-level temporal feature learning [30]. For spatio-temporal data, our research integrates GNNs with recurrent neural networks (RNNs) and long short-term memory (LSTM) networks to mitigate information loss from separate modeling approaches [31]. By integrating multimodal information, including text, images, and waveform signals, into a unified framework comprising spatial nodes and time series, this approach extracts cross-modal spatial features using GNNs. It integrates temporal feature relationships using RNNs and LSTMs, achieving the extraction and fusion of spatio-temporal features through joint training. The specific technical approach is illustrated in Figure 1.

4.2 Construction of a multimodal fusion model and its interpretability analysis

4.2.1 Feature-level fusion of multimodal data Given the multi-source, heterogeneous, and spatio-temporal nature of objective clinical data in TCM during the window period, this study employs a combined feature-level and decision-level fusion approach. By integrating multi-source information at both the feature and decision levels, the study achieves more comprehensive and accurate analysis and decision-making [32, 33]. At the feature level, this study proposes using cross-modal attention to

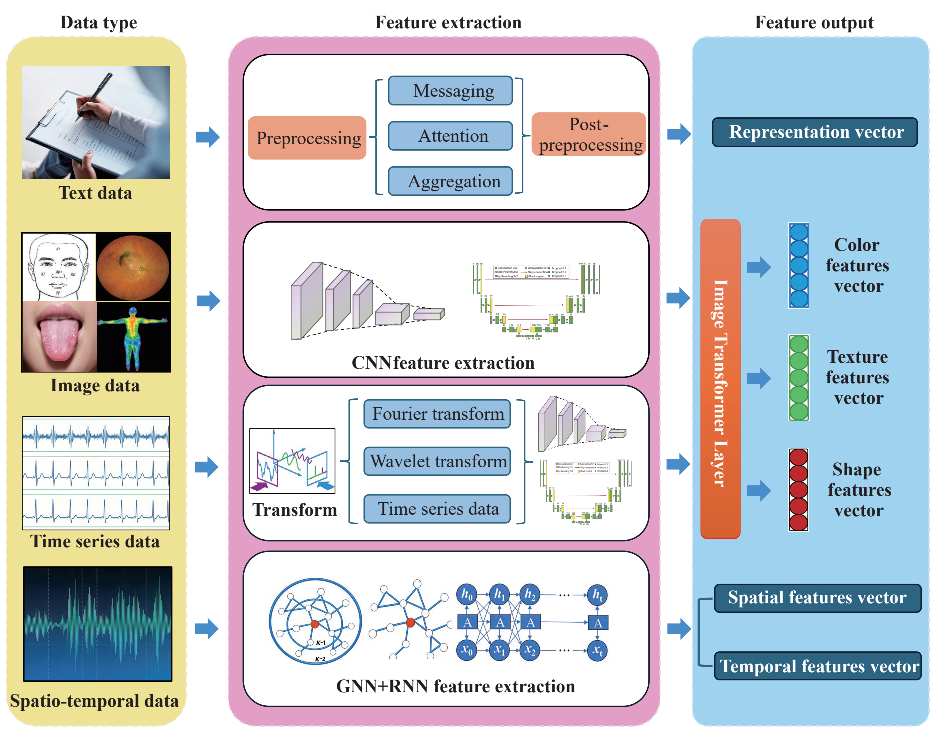

Figure 1 “Macro-meso-micro” three-dimensional multimodal data processing and characterization

fuse features from different data sources and methods, enabling dynamic weight adjustment across modalities. For textual data, Transformer models effectively capture semantic information. For images, CNN integrated with an attention mechanism extracts key visual features [27]. For time-series data, converting it into images enables sharing the feature-extraction network with image modalities. By stacking multiple convolutional layers and an attention mechanism, the network can learn complex, hierarchical feature representations more effectively. For spatiotemporal data, GNNs are first employed to extract spatial structural features, followed by the introduction of gated recurrent units (GRUs) to capture temporal dependencies, ultimately forming fused spatiotemporal feature vectors [34]. After obtaining features from various modalities, we integrate a prior knowledge graph and assign weights to the features of different modalities based on it. In the attention mechanism layer, we fuse the extracted image, text, and spatio-temporal features to obtain a text-spatio-temporal-image attention fusion feature, and assign weights to it via the attention network. The overall technical approach to feature-level fusion is shown in Figure 2.

4.2.2 Decision-level fusion of multimodal data To enhance model interpretability and develop more precise fusion models, decision-level fusion is implemented following the classification of individual modalities. Textual features are extracted using an improved Transformer model that leverages self-attention to capture long-range dependencies in clinical text. Textual features combined with tabular data, enriched with knowledge graph priors, undergo classification via deep learning models embedded with feature-selection layers. For image data such as tongue surface and fundus images, as well as image data derived from time-series data (e.g., pulse waveforms), after a frequency-domain transformation, image features are extracted using methods based on improved MobileNets [35], which are then used to construct deep learning classification models. For spatio-temporal data with spatial structures, an enhanced Transformer model is employed to construct the classifier. By incorporating spatio-temporal position embeddings, the model perceives data across both temporal and spatial dimensions. Furthermore, by stacking multiple layers of self-attention mechanisms and feedforward networks, the model can capture complex dependencies across time steps and spatial units, ultimately producing probability distributions for various TCM state factors. Ultimately, multimodal results are fused through methods such as voting, averaging, or Bayesian approaches, highlighting each modality's role in state assessment. The overall technical approach to decision-level fusion is shown in Figure 3.

4.2.3 Analysis of model interpretability and extraction of key parameters To identify core model parameters and establish interpretable comprehensive assessment indicators, SHapley Additive exPlanations (SHAP) analysis is used to quantify the contribution of each input feature to model outcomes, explaining each prediction locally. Concerning the analysis of TCM condition indicators during the window period, by integrating the TCM theories of the "holistic view" and "condition identification", a mapping relationship is established between data feature vectors (such as blood glucose levels and pulse rate) and TCM condition elements (such as location and nature), then calculating their contributions to achieve consistency among the features. For image data, the Ablation-CAM++ technique can be employed to determine the importance of activation maps for the target class by examining the concept of ablation analysis; that is, by progressively removing or "ablating" certain parts of the input and observing changes in the model's output, understanding the model's dependencies on specific inputs. Furthermore, to better reflect the theoretical implications of TCM in the interpretation of results, TCM-based

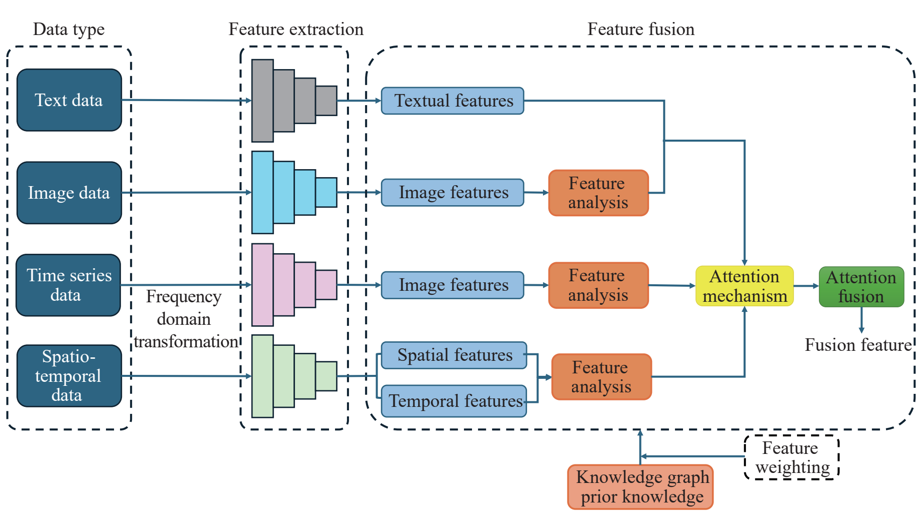

Figure 2 "Macro-meso-micro" three-dimensional multimodal data fusion model at the feature level

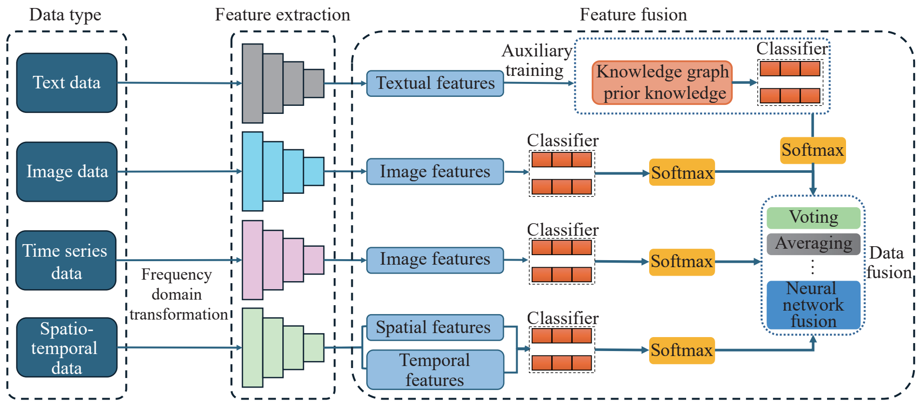

Figure 3 "Macro-meso-micro" three-dimensional multimodal data fusion model at the decision level

functional zoning-guided semantic ablation can be introduced to establish a correspondence between ablation regions and TCM functional zones (e.g., the left cheek corresponds to the liver, the right cheek to the lungs, and the forehead to the heart). In addition, an interpretable image classification framework proposed by PINTELAS et al. [36], integrating segmentation and clustering methods, extracts texture features from image subregions and applies linear white-box predictive models, offering clear and reliable explanations. This facilitates understanding of the relationships between TCM states and various feature indicators, supporting the construction of an evaluation system.

4.3 Development of time-series dynamic efficacy evaluation methods and clinical decision support

In TCM clinical practice, an individual's state elements exhibit highly dynamic evolution over time, particularly during the critical pathological progression of the window period. Traditional static diagnostic methods are inherently limited in representing such dynamic changes, limiting the development and real-time adjustment of individualized interventions. Therefore, establishing a dynamic clinical model that integrates multi-source data to guide clinical decision-making during the window period is important for modernizing TCM. To this end, this study proposes the idea of constructing a diagnostic and treatment model that primarily comprises three components: time-series dynamic representation, information source weight adjustment, and efficacy-feedback assessment.

Initially, building upon preliminary multimodal data fusion, individual time-series data are modeled dynamically using algorithms such as LSTM and GRU [37, 38]. To address non-stationarity and temporal randomness in time-series data, which can cause model overfitting, incremental learning methods are incorporated to enhance real-time adaptability. Given clinical realities, parameter importance may vary by disease stage; thus, information weights require dynamic adjustment over the clinical timeline. To resolve this, attention mechanisms are introduced into models to dynamically adjust multimodal data weights based on prediction errors at different time points, enhancing automatic adaptability to variations in data significance across stages.

In efficacy assessment, causal inference frameworks are introduced, taking current patient health status and intervention strategies as model inputs. By comparing factual and counterfactual model outcomes, the efficacy of individual interventions is evaluated. To improve clinical interpretability of efficacy assessments, semi-supervised learning methods (e.g., support vector machines, co-training algorithms, and generative models) are employed. After labeling a limited subset of the input data, pseudo-labels generated from expert prior knowledge help the model learn from large unlabeled datasets. The overall technical approach to model construction is shown in Figure 4.

Collectively, the proposed model integrates time-series dynamic modeling, adaptive weighting, and efficacy evaluation. By employing incremental and semi-supervised learning strategies, this approach significantly enhances model interpretability, representing a key innovation for individualized TCM clinical decision-making.

4.4 Examples of comprehensive management of prediabetes

Taking the specific "pre-disease to disease" window period of prediabetes as an example, concerning clinical data are collected across three levels: macro level data (including the patient's city of residence, season, and climate), meso level data (including TCM clinical symptoms, four-diagnostic forms, tongue and facial images, pulse patterns, and fundus images), and micro level data (including blood glucose and lipid levels, metabolomics, and gut microbiota), constructing a multidimensional, multimodal dataset for prediabetes. Building on this, the data were classified. For clinical text data related to

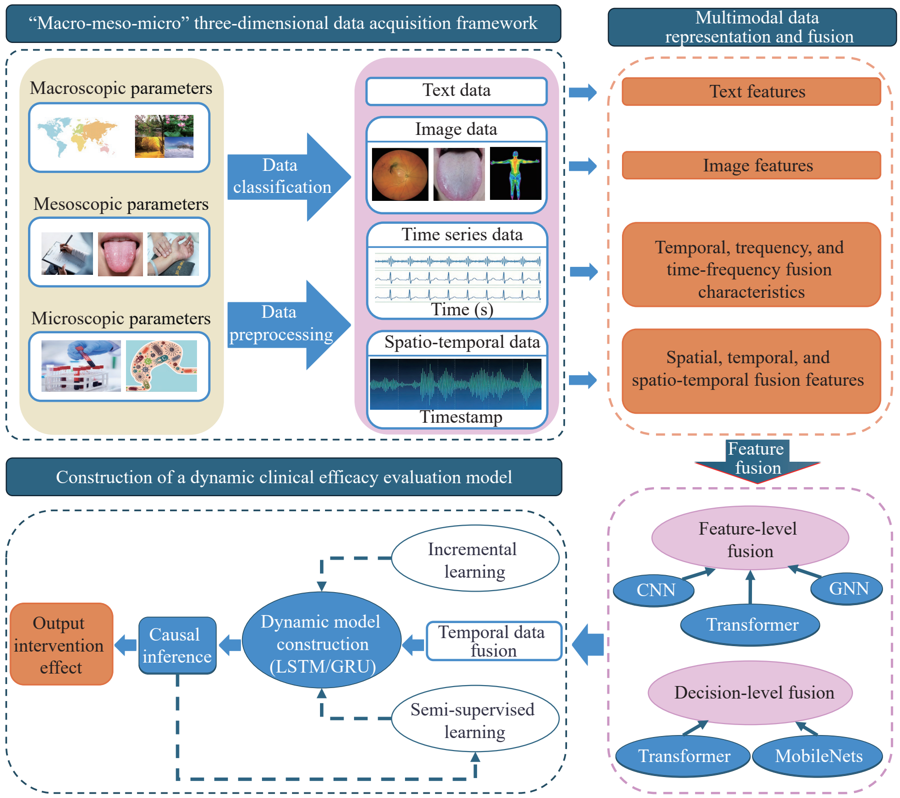

Figure 4 The model for the dynamic assessment of individual clinical management based on time-series data analysis

prediabetes, medical terminology alignment techniques and GNNs were employed to standardize the data representation and extract features. To account for characteristic changes in images of the tongue surface and fundus, a CNN is employed to extract image feature vectors, including color, texture, and position, for joint encoding. For pulse waveform time-series data, sensor parameters, external factors, and temporal effects are taken into account; images are generated via frequency-domain transformation, and a CNN is then used to extract their feature vectors. Finally, taking into account contextual information such as the patient's long-term living environment (e.g., humid and hot regions) and seasonal changes, as well as spatiotemporal information on changes in the distribution of retinal lesions over time, RNN and LSTM were employed to extract temporal features. GNNs were introduced to extract spatial features, enabling the separate extraction of temporal and spatial features.

At the level of feature fusion, for text information related to TCM (such as physical obesity, frequent urination, and thirst), a Transformer model is employed to extract semantic information associated with locational elements (such as the spleen and kidneys) and state-related elements, including Qi deficiency, Yin deficiency, and dampness accumulation. Using a knowledge graph incorporating associations between elements of TCM as prior guidance, the system dynamically adjusts the weights of different modal features to generate a comprehensive feature representation that integrates macro, meso, and micro level data. A CNN employing a fused attention mechanism is used to extract image features, including color, shape, and texture, from tongue and fundus images. Time-series data, such as pulse waveforms and blood glucose levels, are converted into image form for joint feature extraction and fusion with image data. At the decision-fusion level, an improved Transformer network is used to capture associative patterns between symptoms, such as frequent urination and thirst. By integrating prior knowledge of the pathological relationship between "spleen, deficiency, and dampness", a text classifier is constructed that targets factors associated with spleen deficiency and excessive dampness in TCM. Using networks such as MobileNets to extract features, including the thickness of the tongue body, tooth marks, the texture of the tongue coating, the area of hard exudates in the fundus, and the curvature of blood vessels, an image classification model has been constructed for conditions such as spleen deficiency and blood stasis. Time-series data, such as pulse patterns and blood glucose levels, along with spatio-temporal contextual information, such as geographical location and season, are converted into images via a frequency-domain transformation and encoded with spatio-temporal positions. An improved Transformer model is then used to construct a classifier that outputs the probability of each condition. Finally, strategies such as voting, averaging, and Bayesian fusion are employed for back-end fusion, with data weights assigned based on TCM diagnostic logic and clinical prior knowledge, enabling decision support driven by the synergy of multimodal features and knowledge.

Finally, to investigate the transition from “spleen deficiency” to “Yin deficiency” in patients with prediabetes, data, including fasting blood glucose levels, pulse waveforms, fundus images, and daily air temperature, were collected at different time points to construct individual time-series datasets. LSTMs and GRUs were employed to capture evolving trends in data features during the transformation process, whilst incremental learning was introduced to adjust model parameters in stages, adapting to the dynamic progression of prediabetes. At the same time, by dynamically adjusting the weights of the indicators through an attention mechanism, the model’s diagnostic effectiveness for specific disease stages is enhanced: in the early stages, when blood glucose fluctuations are not pronounced but changes in the tongue and pulse are significant, the weights for these indicators were increased; in the progressive stage, when blood glucose fluctuations intensify and microvascular changes in the fundus appear, the weights for blood glucose levels and fundus images were increased. During the therapeutic efficacy feedback phase, a causal inference framework is introduced to quantify the effect sizes of personalized interventions. Combined with semi-supervised learning methods and utilizing labels annotated according to prior rules established by TCM experts, this approach enables the automatic learning of non-linear mapping relationships between therapeutic efficacy indicators, including tongue and pulse characteristics and blood glucose levels, and TCM state factors, providing clinically interpretable feedback on treatment efficacy.

5 Conclusion

Effective clinical diagnosis and treatment during the “pre-disease to disease” window period are crucial for leveraging TCM’s strength in “treating disease before its onset” and achieving early health management. The theoretical advantages of TCM stateology in recognizing critical transition states provide methodological guidance for constructing window period diagnosis and treatment models. This study, centered on state identification, reconstructs the conceptual framework of “symptoms” and proposes a multimodal heterogeneous data representation framework encompassing “macro-meso-micro” parameters. Based on effective multimodal data integration, a dynamic individualized clinical evaluation method rooted in state identification is developed. This approach seeks to address current challenges in precisely identifying and dynamically capturing TCM states during the window period, potentially serving as a reference model to advance intelligent and standardized TCM clinical practice.

Fundings

Key Research Project of the National Key Research and Development Program (2023YFC3503003), Special Project for Central Government’s Guidance on Subnational Science and Technology Development by the Fujian Provincial Department of Science and Technology (2024L3014), and Domestic First-class Discipline Development Project of Hunan University of Chinese Medicine (XJF [2022] No. 57)

Author contributions

Danyang Li: conceptualization and writing – original draft. Min Ai: methodology and writing – review & editing. Pai Zhou: investigation and writing – review & editing. Ying Deng: resources and writing – review & editing. Chaoyang Yang: writing – review & editing. Qinghua Peng: supervision, funding acquisition, and writing – review & editing. All authors approved the submission and take responsibility for this manuscript.

Competing interests

Qinghua Peng is an editorial board member for *Digital Chinese Medicine* and was not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

References

[1] ZHANG QF, HUANG K, SHI XJ, et al. Health management using modern life sciences: exploring the concept of “treating without illness” in the modern context. Scientia Sinica Vitae, 2025, 55(1): 177–184.

[2] 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.

[3] ZHONG LQ, XIN GJ, PENG QH, et al. Deep learning-based recognition of stained tongue coating images. Digital Chinese Medicine, 2024, 7(2): 129–136.

[4] CHEN BB, WANG Y, TANG W, et al. The application ideas of symptomatic treatment from the perspective of traditional Chinese medicine state science. Chinese Journal of Traditional Chinese Medicine, 2025, 66(14): 1439–1443.

[5] World Health Organization. Guidelines on estimation of residual risk of HIV, HBV or HCV infections via cellular blood components and plasma: WHO expert committee on biological standardization. Geneva: WHO, 2018. Available from: https://www.who.int/publications/m/item/annex-4-trs-no-1004.

[6] TANG W, CHAI RT, CHEN H, et al. Research and modern interpretation of the concept of desire for illness. Chinese Medicine Journal, 2024, 65(9): 877–881.

[7] LI CD, ZHAO W, XU JJ. Traditional Chinese medicine health management based on health status identification. Chinese Journal of Traditional Chinese Medicine, 2019, 34(2): 661–664.

[8] ZHU WF. Establishing a new system of syndrome differentiation centered on syndrome element. Journal of Hunan University of Chinese Medicine, 2004(6): 38–39.

[9] LI CD, JI LJ, YANG CY, et al. Construction of a state-centered health cognition theory. Chinese Journal of Traditional Chinese Medicine, 2011, 26(2): 313–316.

[10] MA YM, LIU MR, SONG B, et al. A brief overview of the theory of traditional Chinese medicine's weather-geography-personality integration. Chinese Journal of Traditional Chinese Medicine, 2024, 39(9): 4572–4576.

[11] CHEN BB, WANG Y, TANG W, et al. Summary of application strategies for traditional Chinese diagnostic thinking in the context of undetectable symptoms. Chinese Journal of Traditional Chinese Medicine, 2024, 65(7): 669–673.

[12] CHEN EN, LI SY, HU YQ, et al. Research progress on image acquisition and color information analysis of traditional Chinese medicine tongue diagnosis. Chinese Journal of Traditional Chinese Medicine, 2024, 39(7): 3586–3589.

[13] XIE YY, ZHU L, ZHANG MT, et al. Research on the correlation between the smell pattern of Yin deficiency syndrome in chronic atrophic gastritis and gut microbiota. Chinese Journal of Traditional Chinese Medicine Information, 2025, 32(11): 163–171.

[14] LIU ZH, LUO BB, WU DC, et al. Research on traditional Chinese pulse diagnosis based on three-channel U-shaped micro-nano fiber pulse sensor. Acta Photonica Sinica, 2026, 55(1): 79–91.

[15] QIAO SJ, WU CW, WANG ZL, et al. Thoughts on the development of microscopic diagnosis from the perspective of the basic characteristics of traditional Chinese medicine syndromes. World Chinese Medicine, 2025, 20(23): 4356–4359.

[16] LUO J, ZHANG LN, CHEN JQ, et al. Study on the tongue manifestations of patients with primary Sjögren's syndrome and their correlation with syndromes. Chinese Journal of Traditional Chinese Medicine, 2021, 36(6): 3653–3656.

[17] YANG T, HUANG MN, HE YY, et al. Research on the correlation between TCM syndromes of heart-kidney disconnection in depression and neurotransmitters, and the prediction model for the efficacy of Jiaotai Pill. Chinese Herbal Medicines, 2026, 57(5): 1787–1797.

[18] WANG JQ, ZHANG L, XU SF, et al. Research on the correlation between real-world healthy people's tongue images and age and gender based on machine learning. World Science and Technology-Modernization of Traditional Chinese Medicine, 2024, 26(11): 2806–2814.

[19] KARIM MR, ISLAM T, SHAJALAL M, et al. Explainable AI for bioinformatics: methods, tools and applications. Briefings in Bioinformatics, 2023, 24(5): bbad236.

[20] RONG L, ZOU JY, RAN W, et al. Advancements in the treatment of non-alcoholic fatty liver disease (NAFLD). Frontiers in Endocrinology, 2023, 13: 1087260.

[21] XIAO WD, LI CD, WANG Y, et al. Exploring the issue of "no identifiable symptoms" in chronic stomach and spleen diseases from an overall perspective. Chinese Journal of Traditional Chinese Medicine, 2024, 39(10): 5375–5378.

[22] QIU XR, LENG XJ, LAI XM, et al. Discussion on the optimization methods of the clinical efficacy evaluation system of traditional Chinese medicine from the perspective of individualized diagnosis and treatment. Chinese Journal of Traditional Chinese Medicine, 2024, 65(23): 2389–2393.

[23] QIAN ZZ, DAI XY, ZHANG HQ, et al. Construction of an intelligent decision-making model for traditional Chinese medicine syndrome differentiation and treatment based on multi-modal data fusion: exploration of diagnosis and treatment path optimization driven by deep learning. Chinese Journal of Evidence-Based Medicine, 2026, 26(1): 93–97.

[24] ZHAI HM, LV XJ, HOU ZW, et al. MLNet: a multi-level multimodal named entity recognition architecture. Frontiers in Neurorobotics, 2023, 17: 1181143.

[25] JIN J. A study on knowledge representation learning methods incorporating entity category information. Beijing: Beijing Jiaotong University, 2020.

[26] CHEN SY, SHEN CC, PARK Y, et al. Graph neural networks powered by encoder embedding for improved node learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2026: 1–11.

[27] HONG JC, CHEN SY, POHODICH A, et al. Convolutional neural network-based differentiation of intraocular inflammatory cells with ultrahigh-resolution OCT. Biomedical Optics Express, 2025, 16(12): 4917.

[28] ANGARA S, TRAN L, KIM J. Glaucoma classification using a NFNet-based deep learning model with a customized hybrid attention mechanism. Diagnostics, 2026, 16(5): 815.

[29] PAZO EE, LIU XY, LIU SK, et al. Validation of the eyerobo FC portable fundus camera for diabetic retinopathy screening using public datasets and deep learning. Ophthalmology and Therapy, 2026, 15(4): 1439–1461.

[30] QU S, WU XC, HUANG LG, et al. Diagnosis of disorders of consciousness using nonlinear feature derived EEG topographic maps via deep learning. Scientific Reports, 2026, 16: 7417.

[31] ALASIRY A, SAMPEDRO GA, ALMADHOR A, et al. Epileptic seizures diagnosis and prognosis from EEG signals using heterogeneous graph neural network. PeerJ Computer Science, 2025, 11: e2765.

[32] PANG SC, BI HR, WANG SD, et al. Multi2Fusion: a multiomics fusion framework with multilevel information integration for cancer subtype classification. Journal of Chemical Information and Modeling, 2026, 66(7): 4268–4283.

[33] REN XY, XIU YX, CHEN T, et al. TPVNet: a domain-aware graph-based framework for reliable multivariate physiological time series classification in healthcare. Computer Methods and Programs in Biomedicine, 2026, 276: 109214.

[34] LIU SY, CAO LX. Dynamic spatiotemporal graph attention networks for cross-regional multi-disease forecasting and intervention optimization. Frontiers in Public Health, 2026, 14: 1720620.

[35] OUERHANI A, HADIDI T, SAHLI H, et al. Automatic childhood pneumonia diagnosis based on multi-model feature fusion using Chi-square feature selection. Journal of Imaging, 2026, 12(2): 81.

[36] PINTELAS E, LIASKOS M, LIVIERIS IE, et al. Explainable machine learning framework for image classification problems: case study on glioma cancer prediction. Journal of Imaging, 2020, 6(6): 37.

[37] AVERSANO L, IAMMARINO M, MADAU A, et al. Time series forecasting for bug resolution using machine learning and deep learning models. Frontiers in Big Data, 2025, 8: 1745751.

[38] ESPEJEL-RIVERA MA, TOXQUI-QUITL C, PADILLA-VIVANCO A, et al. Dynamic thermography-based early breast cancer detection using multivariate time series. Sensors, 2025, 25(24): 7649.

(Editor-in-Charge Siyi Wei)

“未病-已病”窗口期的中医临床诊疗模式构建：中医状态学视阈下客观化多模态数据的融合

李丹阳[a], 艾民[b], 周派[b], 邓颖[b], 杨朝阳[a], 彭清华[b, c*]

a. 福建中医药大学中医学院, 福建 福州 350122, 中国
b. 湖南中医药大学中医学院, 湖南 长沙 410208, 中国
c. 湖南中医药大学第一附属医院眼科, 湖南 长沙 410007, 中国

【摘要】“治未病”思想是中医学的基本理念之一，其贯穿中医诊疗体系始终，并发挥着至关重要的作用。然而现阶段“未病-已病”窗口期的中医诊疗模式存在临床参数采集维度缺失、多模态异构数据难以表征融合、干预及疗效评价动态精准性不足等问题。基于此，本文以李灿东教授中医状态学理论为方法指引，聚焦客观化多模态数据融合，提出一种面向“未病-已病”窗口期的中医个体化诊疗新模式。该模式首先提出重构“证”的认知体系，整合宏观、中观及微观等多源异构数据，形成三维评测指标体系的思路；通过引入图神经网络、卷积神经网络、注意力机制及知识图谱引导的权重分配，实现多源数据的协同表征、对齐与融合。其次，规划构建特征级与决策级的多模态融合模型，以建立指标与中医状态要素的映射关系，筛选窗口期病理演变关键指标；提出利用夏普利加性解释（SHAP）、Ablation-CAM++等方法增强模型可解释性的技术路径。最后，以状态评估为落脚点，提出基于长短期记忆（LSTM）网络、门控循环单元（GRUs）等算法，构建基于时序数据分析的个体诊疗动态评估方法的构想，并引入因果推断框架与半监督学习策略，实现个体干预效应量化评估与可解释的疗效反馈，形成从数据表征与融合、权重调整、可解释性分析到动态诊疗反馈的完整技术路径。旨在弥补现有中医“未病-已病”窗口期诊疗模式的不足，为中医“治未病”临床实践提供可操作的方法论遵循。

【关键词】中医状态学；窗口期；临床诊疗模式；多模态数据融合；深度学习