

Applications of Artificial Intelligence in Traditional Chinese Medicine

Xinyi Zhao *

Faculty of Chinese Medicine, Macau University of Science and Technology, Macau, 999078, China

* Corresponding author Email: iriiiis0506@gmail.com

Abstract: Traditional Chinese medicine (TCM) represents a valuable part of Chinese cultural heritage. It is characterized by complex compositions and multiple therapeutic targets, which pose significant challenges to its modernization research. In recent years, the rapid advancement of artificial intelligence (AI) has enabled flexible, data-driven approaches that are now widely applied in TCM studies, including active ingredient analysis, quality control, and safety evaluation. This article reviews recent applications of AI in the TCM field, focusing on several key aspects: the advantages of graph neural network-based network pharmacology in analyzing multi-target interactions; computer vision and hyperspectral imaging for automatic medicinal material identification and quality grading; multimodal fusion techniques for component content detection and production process quality control; and graph attention networks and heterogeneous network representation learning for predicting TCM toxicity and assessing potential risks associated with combined use of Chinese and Western medicines. Based on a narrative review of the latest literature, this paper identifies current challenges such as data standardization and the lack of interpretability in model computation processes. Finally, it offers perspectives on how AI can support the standardization, precision, and global promotion of TCM.

Keywords: Artificial Intelligence; Traditional Chinese Medicine; Quality Control; Graph Neural Networks; Toxicity Prediction.

1. Introduction

The theoretical system of traditional Chinese medicine is based on the two philosophical foundations of 'holistic concept' and 'syndrome differentiation and treatment'. After a long period of practical medication accumulation, it has demonstrated unique advantages in the treatment of chronic diseases, complex diseases and infectious diseases [1]. However, the multi-component and complex composition of traditional Chinese medicine prescriptions involves a wide range of targets in terms of molecular mechanisms, and there are complex biological pathways that work together to make multi-target synergistic mechanisms and individualized diagnosis and treatment models. This makes it difficult to accurately explain the principle of action of traditional Chinese medicine in the way of 'one drug corresponds to one target' [2]. Traditional diagnostic methods rely on 'Inspection, Auscultation and Olfaction, Inquiry, and Palpation' and research and development models are mostly based on experience. These characteristics lead to three major problems in the modernization of traditional Chinese medicine: unclear material basis of efficacy, difficult to explain the mechanism of action, and inconsistent quality control standards. These issues have seriously affected the international recognition of traditional Chinese medicine and its clinical promotion and use [3].

In recent years, large volumes of experimental data have been generated in the field of life science, and artificial intelligence technology is also developing. These two fields are gradually combining with each other. Artificial intelligence algorithms represented by deep learning and large language models are helping to solve some complex scientific problems in the field of traditional Chinese medicine [4]. Compared with the traditional research mode of putting forward hypotheses first and then verifying them,

artificial intelligence technology has high-intensity computing ability, which can not only extract features from a large amount of data, but also deal with complex nonlinear relationships. By integrating multi-omics data, spectral imaging information and ancient books without fixed format, it provides an unprecedented research method to discover the hidden mechanism of action and various influential effects of traditional Chinese medicine [5]. However, most of the existing studies only focus on a single link, such as predicting components or identifying targets, and lack a whole process system from discovering efficacy to evaluating drugs.

2. The Integration of Traditional Chinese Medicine Big Data and Artificial Intelligence Technology

Although the traditional statistical analysis method has been applied in the data analysis of traditional Chinese medicine, there are still many practical problems in the deep integration of artificial intelligence technology into the core research links such as drug property prediction, compound compatibility and target recognition. At present, the modern artificial intelligence Chinese medicine research system that people are building mainly includes three parts, namely, database layer, algorithm layer and application layer.

At the database layer, researchers use natural language processing and knowledge graph technology to perform semantic association and knowledge extraction on ancient literature and large databases such as TCMBank, ETCM, and SymMap, and then construct a multi-dimensional association network containing 'Chinese medicine-component-target-disease' [3]. However, the current TCM data still faces bottlenecks such as insufficient standardization, low degree of data structure, and difficulty in interdisciplinary integration: literature data from different sources are highly

heterogeneous and inconsistent, clinical data are scattered in various institutions and lack a unified recording format and sharing platform[5]. In addition, the existing knowledge map is still relatively rough in expressing the unique concepts of traditional Chinese medicine, such as the concepts of medicinal properties, meridian tropism and syndromes, which is difficult to fully reflect the compatibility law of 'monarch, minister, assistant and messenger' and the dynamic change process of syndrome differentiation and treatment [6].

In the aspect of core algorithm layer, graph neural network can encode the characteristics of network nodes and transmit information through neighborhood diffusion. It can accurately describe the complex topology and interaction between drug molecules and targets, and is particularly suitable for identifying key targets in data-scarce biological networks. [1]. The large-scale language model has shown great potential in clinical knowledge question answering and helping dialectical reasoning. Through the combination of LoRA fine-tuning, prompt engineering and retrieval enhancement generation, it can summarize clinical syndromes and pharmacological laws from real-world data[7]. In the task of identifying phenotypes and evaluating quality, deep learning architectures such as convolutional neural networks and Vision Transformer have performed well in identifying and classifying medicinal image and spectral data, while traditional machine learning methods such as support vector machines and random forests still have their own advantages in the case of a relatively small number of samples[8-9].

3. The Application of Artificial Intelligence Technology in The Field of Traditional Chinese Medicine

3.1. Intelligent Analysis of Pharmacological Mechanism

3.1.1. Multi-Target Synergistic Analysis Based on Network Pharmacology and Graph Neural Network

Network pharmacology is now widely applied to elucidate the holistic pharmacological mechanisms of traditional Chinese medicine. However, conventional network pharmacology typically depends on static physical interaction networks, often overlooking the affinity between drug components and targets as well as the competitive relationships among them. This oversight can lead to numerous false positives in predictive outcomes. By incorporating graph neural networks and deep learning techniques, artificial intelligence is capable of uncovering nonlinear topological associations within large-scale, data-sparse biological networks. AI particularly machine learning, deep learning, and graph neural networks, has significantly advanced network pharmacology, giving rise to methodological framework known as "AI-network pharmacology." This framework enables systematic and precise analysis of the action mechanisms of traditional Chinese medicine, spanning from molecular interactions to patient-level therapeutic effects[1].

Within the theoretical context of the "network target" concept, Li's team further integrated network-target-based computational and experimental approaches with artificial intelligence and multimodal multi-omics technologies, markedly enhancing both the accuracy and scope of network

pharmacology methodologies. This integrated system covers diverse aspects, including predicting disease molecular mechanisms, inferring drug-target relationships, analyzing the compatibility rules of traditional Chinese medicine formulas, and assessing drug toxicity[2]. Fan Xiaohui's team systematically outlined four core steps in network pharmacology research namely, component identification, network construction, network analysis, and experimental validation and emphasized that artificial intelligence can play a significant role at each of these stages[10]. Chen Yuqian's team proposed a new analytical framework at the theoretical level of traditional Chinese medicine prescription design, distinguished the two prescription construction paradigms of 'top-down' and 'bottom-up', and introduced network elements such as high contribution nodes, low contribution nodes and intermediate nodes to transform the prescription optimization problem into a quantifiable network analysis task[11].

In terms of algorithm and platform development, the POINT network pharmacology platform developed by Zhao Yi's team integrates 1,976 multi-level biological networks, covering single cell, tissue and disease-specific networks, thus breaking through the limitations of single PPI network in traditional research. The platform combines the restart random walk algorithm with deep learning methods such as BACPI to avoid the analysis deviation caused by the ranking of node degree values. At the same time, it integrates classical algorithms such as restart random walk and deep learning methods such as BACPI, and constructs a biomedical knowledge graph containing more than 860,000 entities and nearly 6 million relationships, so that the coverage of prediction tasks reaches more than 90%[12]. In a long review published in 'Chemical Reviews' by Hou Tingjun and Xie Changyu's joint team, the application progress of GNN in molecular property prediction, virtual screening, molecular generation, knowledge graph construction and synthetic path planning was systematically summarized, and the integration path of frontier paradigms such as geometric GNN, self-supervised learning and meta-learning with GNN was emphatically discussed, which provided a systematic theoretical support for the innovation of GNN method in network pharmacology. The AnnoNet-TCMs model proposed by Lu Jiang's team is a highly interpretable multi-dimensional heterogeneous graph neural network framework, which integrates five types of entities : herbs, compounds, protein targets, gene ontology entries and disease information. The model achieved a mAUC of 0.8235 in the compound-disease association prediction task, and the recall rate of the first 20 % in the grass drug-disease association prediction task was 0.5706; At the same time, the synergistic combination of tanshinone I and astragaloside IV, a previously unreported myocardial infarction, was successfully identified, and its mechanism of action was verified by experiments involving apoptotic pathways and NF- κ B signaling cascades [13]. This work provides a general framework for interpretable drug combination discovery.

From the perspective of horizontal comparison, each team shows a clear division of labor and complementary relationship in the above research. At the level of theoretical contribution, Li's team's network target theory has source significance. Chen's team's prescription network design paradigm has achieved a theoretical breakthrough in the field of prescription optimization, and Fan's team has contributed a set of standardized four-step operation process. In terms of algorithm focus, Hou Tingjun's team focused on the

construction of GNN basic methods (such as geometric GNN, self-supervised learning), Zhao Yi's team emphasized the systematic integration of multi-algorithm collaboration (restarting random walk and BACPI combination) and multi-omics networks, and Lu Jiang's team focused on the relationship learning and semantic enhancement in heterogeneous graph neural networks. In terms of application scenario coverage, Zhao Yi's team's POINT platform has the strongest versatility and has opened services to users; Li's team covers the whole chain verification from molecular mechanism to patient level; the model of Lu Jiang's team is outstanding in the identification of collaborative new discoveries. In the face of common problems such as sparse data, insufficient network specificity and weak interpretability of prediction, each team adopted a differentiated strategy: Zhao Yi improved the adaptability of the model to environmental information by introducing single cell, tissue and disease-specific networks; Lu Jiang used multi-dimensional heterogeneous graph structure to realize drug prediction and new mechanism discovery under sparse data conditions. Li uses large model pre-training (such as Geneformer and scGPT) to enhance the stability in small sample scenarios.

At present, the field of AI network pharmacology has formed a mature pattern of multi-point breakthroughs and coordinated development. These studies together constitute a complete academic chain from theoretical construction, algorithm development, platform development to clinical verification. In the future, further promoting the cross-integration and benchmark testing of representative methods of each team will help promote the development of network pharmacology to a more accurate, more interpretable and closer to clinical transformation.

3.1.2. Graph Neural Network and Multi-Target Synergy Mechanism Quantification

Graph neural network (GNN) can encode the characteristics of network nodes and transmit information through neighborhood diffusion. It can find nonlinear topological correlations in biological networks with large scale and sparse data, accurately calculate the core targets that determine the efficacy of traditional Chinese medicine, and also quantify the synergistic relationship between different active molecules. Recently, an important research work has constructed an interpretable graph artificial intelligence (Graph AI) framework. The researchers constructed a multi-dimensional knowledge graph (TCM-MKG) containing seven standardized modules of traditional Chinese medicine terms, Chinese patent medicines, Chinese herbal pieces, pharmacological sources, compounds, biological targets and diseases, and then used the neighborhood diffusion strategy to increase the coverage of the association between compounds and targets from 12.0% to 98.7%. This framework also adds the 'nature, taste and meridian' attribute of the drug as a virtual node, so that the graph neural network can quantitatively describe the classic compatibility roles of 'monarch, minister, assistant and messenger', and can also reversely deduce the pathogenesis type of traditional Chinese medicine from the model of diagnosis and curative effect [6].

In the field of multimodal deep learning, the MDL-HTI framework combines heterogeneous graph learning with biological data coding. These biological data include drug components, target pathways, and ligand properties. It uses multi-view heterogeneous relational embedding algorithms to extract the structural patterns of subgraphs. It also introduces

a self-attention mechanism to achieve dynamic fusion of different features, and performs better than existing benchmark methods in predicting drug-target interactions[14]. In the aspect of drug-target interaction prediction model, DTI-BGCGCN innovatively combines the target attribute graph of bipartite graph drug with the clustering graph convolution network. It can classify different graph topologies, so as to speed up the training. At the same time, this model is suitable for both modern drug data sets and traditional Chinese medicine data sets, which provides an efficient and accurate technical path for predicting drug-target interaction[15].

3.1.3. Cross-scale Pharmacological Analysis Framework

Hong et al. proposed an artificial intelligence-driven multi-scale target analysis framework. This framework can expand the research of traditional Chinese medicine from the pharmacological effect of a single compound to the synergistic mechanism of a single herb and a compound [16]. This framework has several core ideas. On the one hand, it will take the monomer target predicted at the molecular level as the basic input content, and use these contents to analyze the synergistic network between multiple components within a single medicinal material. On the other hand, it will also integrate these modular networks to reveal the complex multi-target synergy mechanism at the level of traditional Chinese medicine compound. This can lay the foundation for the intelligent recommendation of the compound and the accurate prediction of the dosage[16]. In addition, this framework not only provides a bottom-up integration path, but also explicitly advocates a closed-loop verification model called 'calculation-experiment'. This model can help bridge the "semantic gap" between abstract TCM theory and molecular data, and then promote TCM research to the fifth paradigm [16].

At the molecular level, this framework emphasizes the application value of artificial intelligence in identifying single compounds and their direct targets. Researchers can use deep learning models, such as graph neural networks or attention mechanisms, to learn features from chemical structure and biological activity data. In this way, the rapid prediction of a large number of natural product molecular targets can be achieved, and the prediction cost is relatively low[16]. These prediction results can provide a reliable node and edge basis for the construction of subsequent multi-component networks.

At the level of single herb scale, this framework further focuses on the synergistic relationship between different compounds within the same herb. Researchers will take the monomer target and its interaction relationship obtained on the molecular scale as the basic unit, and then use the network propagation algorithm or module detection algorithm, such as restarting the random walk algorithm or community discovery algorithm, to construct a multi-component collaborative network. The purpose of this process is to identify those biological pathways or disease-related sub-network modules that multiple compounds within the same medicinal material jointly target. In this way, it can be clarified that the different components in a single herb do not work in isolation, but rather achieve synergy through functional modularization. It is an important development direction of modern network pharmacology methodology to extend the analysis object of network pharmacology from single compound-target association to multi-component synergistic relationship.

At the level of compound scale, this framework will construct an integrated multi-level network by aligning and

integrating modular networks from different single herbs across herbs. This operation will map the specific functional modules of each medicinal material to the same disease network framework, so as to identify the synergistic clusters between multiple medicinal materials and multiple targets. This cross-scale network integration strategy can reveal the overall regulatory mechanism of the compound, and also explain why the compound has better clinical efficacy than a single drug or a single component at the system level. The 'computational-experimental' closed loop advocated by this framework will ensure the reliability of cross-scale inference through two aspects of work.

One aspect is to compare the predicted results with the known compatibility data, such as the compatibility records of drug pairs, to see how consistent they are. Another aspect is to perform perturbation verification in cell models or animal models, for example, by knocking out some key predictive targets for verification[16]. This systematic verification strategy can provide a solid biological support for the results predicted by artificial intelligence. This study provides an important methodological basis for the systematic analysis of the pharmacodynamic mechanism of traditional Chinese medicine and the precise medication driven by artificial intelligence [16].

In the aspect of synergistic compatibility of natural products, artificial intelligence will adopt the method of identifying natural product drug combinations to integrate various types of data from different data sources. These data include chemical structure data, multi-omics information, and pharmacological network data. Artificial intelligence can also automatically extract features from these data. In this way, natural product pairs with synergistic effects can be quickly screened out in a large number of combination spaces, thereby significantly improving the accuracy of prediction. The synergy of natural products can be achieved through a variety of mechanisms, mainly in these three categories. The first type is pharmacokinetic synergy, in which one compound enhances the absorption effect or bioavailability of another compound. The second category is pharmacodynamic synergy, that is, different compounds act on different targets in the same disease pathway, resulting in a superimposed amplification effect. The third type is network-level coordination, which is to regulate parallel or compensatory signal modules to overcome the redundancy problem on the pathway[17].

Traditional high-throughput screening methods and conventional calculation methods are often affected by the fragmentation of experimental data, high cost and huge combination space, so it is difficult to effectively promote the discovery of natural product drug combinations. In order to break through these bottlenecks, researchers have developed a variety of artificial intelligence models [18]. For example, deep neural network models such as DeepNPD and DeepCombo, which use HERB database and DrugCombDB database respectively, can successfully predict drug combinations in natural products and synergistic effects of chemotherapeutic drugs[19]. For example, the DeepMDS model was used to predict the combination of synergistic anti-breast cancer compounds in the herb pair of *Prunella vulgaris*-*Dandelion*, and then the prediction results were verified by liquid chromatography-mass spectrometry and in vitro cell experiments. Multiple multi-component synergistic combinations such as F973, T271 and T1685 were successfully identified, confirming that these combinations

did have significant synergistic anti-tumor activity at the actual concentration ratio of the extract. Preliminary bioinformatics analysis showed that these compounds could synergistically regulate tumor-related pathways, inhibit the growth of tumor cells, induce apoptosis, and block the process of cell cycle [20]. In the direction of multimodal data fusion, the MDL-HTI framework integrates heterogeneous graph learning and multimodal biological data, and uses a multi-view heterogeneous relation embedding algorithm to extract subgraphs, meta-paths, and structural patterns in the community. In addition, the MAMGN-HT model constructs a heterogeneous graph containing herbs, efficacy, components and targets. By integrating the meta-path and attention mechanism, the accuracy and interpretability of TCM target prediction are improved[19]. MRHormer model constructs a Chinese medicine-target heterogeneous information network, which introduces topological robust local attention mechanism to alleviate the hidden bias caused by the difference of type features, and also uses multi-relational global attention to capture high-order features and long-range dependencies. The most advanced prediction performance is achieved under the inductive link prediction setting [21]. At the same time, the artificial intelligence method based on molecular compatibility theory will mine the synergistic mode from multi-omics data and structural data, which provides a systematic framework for the modeling and optimization of Chinese medicine compound combination. On the basis of these works, the deep integration of artificial intelligence and multi-omics technology is driving the research and development of natural medicines towards precision innovation. This integration will build a network model of 'component-target-disease', and integrate multi-omics technology to form a complete 'virtual prediction-experimental verification-mechanism analysis' research process[17].

These studies together illustrate one thing : artificial intelligence-driven multi-scale analysis frameworks and collaborative prediction methods for natural products are driving the research of traditional Chinese medicine from the traditional single-target model to the direction of systematization and precision. In recent years, the systematic combination of artificial intelligence and network pharmacology, referred to as AI-NP methodology, has effectively overcome the key bottlenecks encountered by traditional network pharmacology in multi-dimensional data integration and heterogeneous data association analysis through advanced feature extraction technology and deep representation learning technology, thus constructing a multi-level drug-target-disease network with biological significance [1].

3.2. Automated Evaluation of Quality Control of Chinese Medicinal Materials

3.2.1. Appearance Identification of Medicinal Materials Based on Computer Vision

The appearance characteristics of Chinese medicinal materials, such as shape, color and section texture, are the basic basis of traditional quality evaluation. However, this evaluation method is largely inseparable from experienced manual judgment. With the rapid development of artificial intelligence technology, computer vision technology based on convolutional neural network and Vision Transformer has been able to realize the automatic processing of medicinal materials identification through transfer learning and deep

feature extraction. At the same time, this technology has made a lot of progress in image recognition, odor analysis, authenticity identification, origin traceability, quality grade division and storage life determination [22].

In terms of model development and performance, several teams have carried out systematic research. The research team of Macao University of Science and Technology has constructed a high-quality database containing more than 150,000 high-definition Chinese herbal medicine images, which is one of the high-quality Chinese herbal medicine image databases in China. Wang et al. collected 276,767 images covering 163 kinds of Chinese medicinal materials, and compared the five models of EfficientNetB0, ResNet50, MobileNetV3, VGG19 and ResNet18. It was found that the accuracy of EfficientNetB0 on the validation set reached 99.0%, and the AUC reached 0.9942. On the test set, the accuracy was 99.0%, the sensitivity was 99.0%, the specificity was 100.0%, and the AUC was 1.0. Aiming at the YOLO series target detection algorithm, some studies have collected images of 148 kinds of Chinese herbal pieces, improved the YOLOv3 algorithm, and introduced CIoU loss function and DIOUNMS. The average detection accuracy on the test set reached 98.47 %, which was 1.83 % higher than the original YOLOv3 [23]. In addition, AAFM-YOLO model is proposed and tested on the self-built TCM-Herbs2025 dataset. The results show that the model shows good robustness in complex recognition scenarios[22].

In the application of Vision Transformer, a data set was constructed based on 611 batches of Angelica Dahuricae Radix decoction pieces, and three models of SVM, CNN and Transformer were used to classify the quality grades. The results show that the Transformer model performs best, with an accuracy rate of 95.1%, which is better than CNN's 91.8% and SVM's 85.2%. The interpretability analysis of SHAP further showed that the color feature contributed the most to the quality classification, and the color mean value was a key discriminant index[24]. ResNet and ViT were also used to classify the images of different processing grades of *Eucommia ulmoides*. The results showed that the accuracy of ViT model was more than 95 %[25]. On the international public dataset, the BEiT model achieved an accuracy of 99.14 % on the VNPlant-200 medicinal plant database[26].

In terms of specific application scenarios, authenticity identification is one of the typical application directions. There is a study on the authenticity identification of *Cordyceps sinensis*. Five models, LeNet-5, AlexNet, GoogleNet, ResNet50 and VGG16, are compared. The results show that VGG16 performs best. The accuracy of the training set is 100%, the accuracy of the validation set is 96.97%, and the macro average F1-Score of the validation set is 94%. From the confusion matrix, the recall rate of pseudo cordyceps reached 100 %, and the recall rate of true cordyceps was 88 %[27].

In terms of origin traceability, there are also many research cases based on the combination of spectral technology and deep learning. In a study, near-infrared spectroscopy combined with Dual-TCNet model was used to identify the origin of *Astragalus membranaceus*. The accuracy of 98.31% was achieved on 588 *Astragalus membranaceus* samples from five producing areas, and the accuracy rate was 98.39%. The recall rate was 98.32%[28]. Another study used terahertz spectroscopy and CNN-Transformer hybrid network algorithm to identify the origin of *Pueraria lobata*, and achieved an accuracy of 91.67% on samples from eight

producing areas. The F1-Score of Jiangxi and Shaanxi samples reached 1.000[29]. Aiming at the origin traceability of *Salvia miltiorrhiza*, a method combining hyperspectral imaging and improved DeiT model was proposed. The image color and short-wave infrared spectral data were fused, and the classification accuracy reached 98.72%[30].

In terms of the origin traceability of *Angelica sinensis*, some studies were based on the data of 16 functional compounds and 40 trace elements, and combined with machine learning algorithms to classify samples from eight different origins. The prediction accuracy of the training set and the test set reached 100.00%[31].

In terms of quality grade classification and processing process evaluation of decoction pieces, a set of system for quality grade classification of *Angelica Dahuricae* decoction pieces is relatively mature. A qualitative discriminant analysis of different processed products of *Radix Ophiopogonis* showed that the method of near-infrared spectroscopy combined with convolutional neural network could achieve rapid non-destructive detection and effectively improve the reliability of discrimination[32]. From these cases, we can see a clear rule : in the tasks of authenticity discrimination, origin traceability and grade classification of ginseng, angelica, sanqi and other bulk medicinal materials, models based on computer vision and deep learning can achieve relatively high accuracy, basically above 90%. These models can automatically eliminate interference under complex illumination or different shooting angles. The research data show that in the image recognition task, the artificial intelligence model can effectively capture some subtle texture differences that are not easy to detect by the naked eye. It is better than the traditional manual identification in terms of recognition efficiency and objectivity, and provides a strong technical support for improving the efficiency of quality control of Chinese medicinal materials[22].

3.2.2. Nondestructive Testing of Hyperspectral Imaging and Spectral Fusion Technology

Hyperspectral imaging technology combines the advantages of spectral analysis and machine vision, which can simultaneously obtain the spatial information and spectral information of the sample [33]. In the study of the origin traceability of *Salvia miltiorrhiza*, the researchers proposed a new framework that combines the two-dimensional correlation spectral image and the improved data-efficient image conversion model. They enhanced the model's ability to focus on key bands by introducing a convolutional block attention module, and achieved 99.62% accuracy of the test set using only 79 wavelengths. The classification basis of the model was also visually verified by hierarchical class activation mapping, which proved the interpretability and reliability of the deep model in origin traceability[33].

In terms of multimodal fusion, researchers use deep learning to fuse the spectral features and color features of hyperspectral imaging, which provides a more reliable discriminant scheme for the authenticity identification and origin traceability of genuine medicinal materials such as *Salvia miltiorrhiza* [34]. In addition, in the non-destructive identification of *Pseudostellariae Radix*, the researchers systematically compared the performance of machine vision and short-wave infrared hyperspectral imaging, and further identified the characteristic compounds by ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry, and established a 'imaging-spectroscopy-mass spectrometry' trinity non-destructive

identification framework [35]. These studies have shown that the integration of hyperspectral imaging technology with deep learning and machine learning models is promoting the quality control of traditional Chinese medicine from 'relying on empirical judgment' to 'digital measurement'.

3.2.3. Quantitative Analysis of Spectroscopic Techniques and Chemometrics

Traditional chromatographic techniques such as high-performance liquid chromatography and gas chromatography-mass spectrometry, despite their high accuracy, are often limited by long detection times, sample damage, and complex operational procedures. In contrast, modern spectral analysis methods, including near-infrared and Raman spectroscopy, when combined with chemometrics and artificial intelligence approaches, offer intelligent solutions for quality control in traditional Chinese medicine manufacturing processes [36]. These methods not only enable traceability of medicinal materials, detection of adulteration, species identification, and differentiation between wild and cultivated varieties, but also cover full-process quality monitoring in TCM production. This includes real-time determination and monitoring of intermediate components, chemical profiling and biomarker evaluation, as well as analysis of dissolution behavior and shelf-life determination of formulations[36].

3.2.4. Comprehensive Evaluation System of Multimodal Fusion

As TCM quality is influenced by multiple factors: such as origin, processing technology, growth environment, harvesting time, it is difficult to fully reflect its internal quality with only one analysis method. The current research trend is to construct a multi-modal fusion evaluation system, using the attention mechanism or feature fusion strategy of artificial intelligence to systematically integrate different types of data such as trait images, spectral components, chemical fingerprints and biological potency. This cross-dimensional modeling method can ensure a consistent assessment of the entire process of medicinal materials from the source to the drug. A comprehensive review pointed out that the artificial intelligence quality control system should further integrate multi-omics and bioinformatics methods, such as network pharmacology, so as to achieve efficacy-oriented quality evaluation, including the discovery of quality markers, elucidation of pharmacodynamic mechanisms, and prediction of endogenous and exogenous toxic metabolites[8].

3.3. Detection and Early Warning of Toxicity Risk of Traditional Chinese Medicine

3.3.1. Toxicity Prediction Based on Heterogeneous Network.

Gao constructed an herb toxicity prediction framework based on heterogeneous networks, called 'HerbToxNet'. This framework first constructs a heterogeneous network containing herbs, components and molecular targets, and then uses the heterogeneous graph attention network to learn the deep features of herbs. Then, the comparative learning mechanism is introduced. By dynamically adjusting the coefficient, the characteristics of herbs with the same toxicity label are drawn closer, and the characteristics of other herbs are pushed away, so as to improve the quality of the characteristics. Finally, the weighted label fusion strategy is used to enhance the prediction results by using the toxicity labels of similar herbs[5].

3.3.2. Prediction of Drug Interactions

The combination of Chinese and Western medicine is a very important issue in clinical practice, but the potential risk of interaction also needs to be paid enough attention [37]. In order to deal with the problem of adverse reaction prediction caused by the complexity of traditional Chinese medicine compound composition, different research teams have developed a variety of artificial intelligence prediction methods.

Yang et al. proposed a dual-graph attention network[37]. This network first extracts the key structural features inside the active molecule, and then captures the spatial structure relationship between different compounds through the attention mechanism. Among advanced deep learning methods such as graph convolutional networks, Weave models, and message passing neural networks, this model performs better in prediction[37].

In terms of incompatibility prediction, Ma et al. designed a deep attention network[38]. This network builds an attention aggregation module to learn the composition characteristics of the drug pair level, and also uses the similarity distribution to characterize the efficacy and drug properties. This model further defines common constraints and specific constraints, and realizes dynamic feature fusion between drug pairs through gated attention units. The model performs well on the IHPTCM dataset and can also provide interpretability analysis at the component level[38].

In addition to the above two models, there are currently cutting-edge methods such as TEmbed-DDI, CWI-DTI, and a prediction framework based on the TCMBank database. The research team of City University of Hong Kong proposed the TEmbed-DDI framework[39]. This framework combines the embedded representation generated by the large language model with the bidirectional neural network, and specifically targets the drug interaction labeling for the combination of Chinese and Western medicines. In the performance test, TEmbed-DDI achieved an AUC value of 0.992 on the Western drug CHCH interaction benchmark and an AUC value of 0.95 on the DEEP interaction benchmark[39]. The CWI-DTI model developed by some scholars is used to predict the interaction between drugs and targets[40]. This model can efficiently capture the similarity between drugs and targets, and has achieved good results in the task of integrating Chinese and Western medicine data sets[40].

In terms of database support, the Chen team of Sun Yat-sen University has built a large-scale TCM database, called TCMBank[41]. This database covers 9192 herbs, 61966 components, 15179 targets and 32529 diseases, as well as their association information. Researchers used the DDI model to perform unsupervised learning based on large-scale data from TCMBank to predict adverse reactions between Chinese and Western medicines. In practice, this method is aimed at the combination of a traditional Chinese medicine and a western medicine. First, all the active components contained in the traditional Chinese medicine are queried through TCMBank, and then whether any of the components may have adverse reactions with western medicine is judged [41].

In terms of clinical verification, some studies have begun to verify based on real-world data. A research team has extracted historical data from the US FDA 's adverse event reporting system (FAERS), and combined with the constructed knowledge map of adverse reactions of Chinese and Western medicines, realized the prediction of adverse

reactions of traditional Chinese medicine formula granules combined with Western medicine[42].

On the whole, the AI prediction of the safety of Chinese and Western medicine is no longer limited to a single model. From the early graph attention network, to a variety of methods that integrate large language models, heterogeneous graph embedding, and deep learning of large-scale databases, researchers are building a multi-level prediction system. These prediction results are also gradually combined with real-world data to provide more reference for clinical safe medication.

3.3.3. Intelligent Monitoring of Clinical Medication Safety

In real clinical scenarios, with the help of natural language processing technology to deeply analyze electronic medical records and adverse reaction reports, artificial intelligence early warning systems can dynamically identify the intricate risk networks behind the combination of Chinese and Western medicine. Combined with the patient's personal age, weight, liver and kidney function and past medical history, the clinical decision support model can make a more accurate prediction of the reasonable dosage. In recent years, many research work in this field has achieved good results.

On the one hand, NLP technology has shown certain advantages in the efficiency of entity extraction of adverse drug reactions. In a study on the real-world safety analysis of Xuebijing injection, researchers used a total of 111758 electronic medical record data from 9 tertiary hospitals across the country, and proposed a comparison framework including 'manual labeling model' and 'retrieval enhanced generation labeling model'. The results show that the RAG labeling model can efficiently extract the entities of Xuebijing adverse drug reactions such as nausea and vomiting by relying on the predefined dictionary and automata adaptation technology. The character-level F1 values are all greater than 96 %, and the labeling efficiency is 5 times higher than that of manual[43]. This shows that in the preprocessing task of large-scale structured data, NLP can indeed complete the named entity recognition task more efficiently, saving a lot of manpower for subsequent adverse reaction analysis and risk network construction.

On the other hand, in terms of the ability to identify the causal relationship of adverse events, the large language model based on Transformer architecture also performs better than traditional methods. In a study based on the Singapore discharge summary, the researchers used the LLM-BiLSTM architecture to identify the causal relationship between unstructured text Chinese medicine and adverse events[44]. The results show that the large language model based on Transformer performs better than the traditional machine learning method, and the F1 score is significantly improved compared with the previous benchmark method based on BioM-ELECTRA-Large-BiLSTM[44]. This shows that the large language model has stronger relationship recognition ability when dealing with the drug-adverse event relationship in unstructured text such as discharge summary.

In terms of molecular toxicity prediction, the large language model represented by GPT-4 also shows good potential. Yang Hengzheng et al. used GPT-4 combined with molecular docking technology to deeply analyze the potential cardiotoxicity of medicinal and edible Chinese herbal medicines[45]. In this study, the performance of GPT-4 was systematically compared with traditional deep learning or machine learning models (such as WeaveGNN, MorganFP-

MLP, SVC, KNN) for different endpoints such as bone toxicity, neurotoxicity, and reproductive toxicity. The results showed that the performance of GPT-4 in some toxicity prediction tasks was not inferior to that of traditional models, and even slightly better in some aspects [45]. The researchers also further combined molecular docking technology to focus on the binding affinity of components in black sesame, ginger, perilla, toosendan, ginger, turmeric, licorice, yam, Indian currant and nutmeg to the cardiac target Cav1.2. It was found that these components had significant binding affinity with Cav1.2, supporting the hypothesis that they may cause cardiotoxicity[45]. It is worth noting that this study shows that as long as the appropriate corpus is established, even without professional computer knowledge or programming skills, high-accuracy learning models can be generated through GPT-4, which gives more researchers without deep computer background access to advanced toxicity prediction tools[45].

However, it should also be noted that the reliability of the large language model in drug safety assessment is not flawless, and it is still necessary to maintain a certain cautious attitude when using it. A study published in 'Drug Discovery Today' systematically evaluated the performance of GPT-4 in the evaluation of organ-specific drug toxicity in liver, heart and kidney, and systematically compared it with the expert evaluation criteria established based on the US FDA drug instructions [46]. The results show that the accuracy of GPT-4 is 48 % to 72 % under the condition of General Prompt (simulating the conversational questioning method of the general public). Under the condition of Expert Prompt (optimized expert questioning method), the accuracy rate is increased to 64 % to 75 %[46]. In general, the performance of GPT-4 is at a medium level, indicating that its ability still has room for further improvement when dealing with more sophisticated drug safety assessment tasks. To this end, the study suggests that in sensitive areas involving public health, if GPT-4 is used to assist decision-making, it may be necessary to use a framework such as retrieval-enhanced generation (RAG) to make full use of the existing knowledge in GPT-4 to further improve its reliability[46].

If the big language model and toxicological genomics data are further integrated in the future, it is expected to establish a complete artificial intelligence-assisted safety system from molecular toxicity mechanism to individualized drug warning. From the molecular level to do fine toxicity prediction, and then combined with the patient's personal characteristics to do individualized medication warning, this system can provide more comprehensive guarantee for clinical safe medication.

4. Conclusion

With the continuous advancement of modern technology, artificial intelligence has achieved in-depth integration with traditional Chinese medicine (TCM) research. This innovative integration effectively breaks through the bottlenecks and constraints inherent in conventional TCM research approaches, freeing TCM research from overreliance on empirical summary and clinical experience alone. Currently, TCM research is gradually transforming into a modern research system dominated by data analysis. Artificial intelligence technology enables systematic exploration of TCM pharmacological mechanisms, multi-dimensional quality evaluation and control of TCM drugs, and early identification of potential safety risks in TCM medication, demonstrating outstanding application value and

broad development prospects in TCM modernization research. Nevertheless, against the backdrop of rapid technological integration, the current AI-assisted TCM research still faces numerous practical challenges that need to be addressed to support the high-quality development of TCM in the future.

One prominent problem in current TCM digital research is the lack of unified and standardized data specifications. TCM clinical and experimental data collected by different medical institutions, research platforms and scientific research teams are fragmented and isolated, failing to form effective data integration and sharing mechanisms. To solve this problem, it is essential to build a professional, standardized multi-modal TCM database covering multiple core data types, including textual knowledge mined from ancient TCM literature, chemical characteristic fingerprints of TCM materials, multi-omics biological data, and real clinical practice data. Meanwhile, unified data metadata standards and comprehensive whole-process data quality management mechanisms need to be formulated and implemented to standardize the collection, sorting and application of TCM data resources.

Existing deep learning models applied in TCM research still have deficiencies in interpretability, and the internal computational logic and derivation process of intelligent algorithms cannot fully match the core connotation of TCM syndrome differentiation and treatment theory in cognitive logic. In future research, advanced artificial intelligence methods such as zero-shot learning, end-to-end model architecture and self-supervised contrastive learning can be adopted to optimize and upgrade the existing algorithm frameworks. Through continuous iterative optimization, the overall stability, generalization ability and practical applicability of intelligent models in TCM scenario applications can be effectively improved.

There is an inherent cognitive gap between the macroscopic and abstract theoretical system of traditional Chinese medicine and the microscopic molecular-level biological data obtained from modern experiments. To bridge this gap and realize the organic combination of TCM theoretical research and modern technological means, the industry needs to construct a closed-loop research mode that integrates intelligent computational analysis and experimental verification. The research process can follow the logical path of intelligent preliminary analysis, experimental verification and model iterative optimization. Specifically, artificial intelligence is firstly used to conduct data mining and preliminary result inference, then in vivo and in vitro biological experiments are carried out to verify the authenticity and scientificity of the preliminary results, and model parameters are dynamically adjusted and optimized based on experimental feedback, so as to guarantee that the computational analysis results are scientifically valid and have practical biological research significance and clinical reference value.

With the continuous innovation and maturity of artificial intelligence technology, the depth and breadth of its integration with traditional Chinese medicine research will be further expanded. The future development of this research field will mainly focus on three core directions: constructing unified and standardized TCM data infrastructure systems, developing interpretable intelligent models that adapt to the theoretical characteristics of traditional Chinese medicine, and building complete research verification systems that connect computational prediction results with actual clinical

effects. The continuous improvement of the above research systems will effectively enhance the repeatability and clinical transformation value of AI-assisted TCM research outcomes, and lay a solid scientific foundation for the standardized, modernized and evidence-based development of traditional Chinese medicine.

References

- [1] Cui, G. Q., Li, M. Z., Guo, W. B., Gao, M., Zhu, Q., & Liao, J. (2025). AI driven network pharmacology: Multi-scale mechanisms of traditional Chinese medicine from molecular to patient analysis. *Computational and Structural Biotechnology Journal*, 27, 5087-5104.
- [2] Wang, Z. Y., Zhang, T. Y., Wang, B. Y., & Li, S. (2025). TCM network pharmacology: new perspective integrating network target with artificial intelligence and multi-modal multi-omics technologies. *Chinese Journal of Natural Medicines*, 23(11).
- [3] Guo, P., Jiang, M., Hu, S., Jiang, Q., Li, L., Wu, J., Ma, Y., & Wu, Z. (2026). Advancing the modernization of traditional Chinese medicine through artificial intelligence and multimodal data integration. *Chinese Medicine*, 21(1), 54. <https://doi.org/10.1186/s13020-025-01194-y>.
- [4] Ren, J., Shi, X., Liang, R., Qin, Y., Gao, X., Zhang, Q., Zhao, L., & Guan, X. (2025). Revolutionizing traditional Chinese medicine research: Advancements in traditional machine learning and deep learning for AI-driven applications. *Journal of Pharmaceutical Analysis*.
- [5] Zhu, Y. Z., Miao, Y. B., Sun, R., Yan, Z. M., & Yu, G. X. (2026). Traditional Chinese medicine toxicity prediction by heterogeneous network. *Expert Systems with Applications*, 299(Part A), 129969.
- [6] Zeng, J., & Jia, X. (2025). Quantifying compatibility mechanisms in traditional Chinese medicine with interpretable graph neural networks. *Journal of Pharmaceutical Analysis*, 15(8), 101342. <https://doi.org/10.1016/j.jpha.2025.101342>.
- [7] Han, C., Yang, G., Li, H., Zhu, L., & Feng, M. (2026). Tuning and clinical application of large language models in Traditional Chinese Medicine: scoping review. *Chinese Medicine*, 21(1), 71. <https://doi.org/10.1186/s13020-026-01346-8>.
- [8] Li, M. Y., Zhu, J. Q., Liu, X. N., Wu, M. Y., Dong, K., Li, X. Y., Gao, P., & Jiang, Z. H. (2025). Application and research progress on artificial intelligence in the quality of Traditional Chinese Medicine. *Frontiers in Pharmacology*, 16, 1687681. <https://doi.org/10.3389/fphar.2025.1687681>.
- [9] Jia, S., Lai, H., Bai, H., Li, M., Zhang, Q., Zhao, P., Wang, Y., Tang, Y., & Leung, E. H. L. (2026). Harnessing the application of artificial intelligence in identification of traditional Chinese medicines. *Journal of Ethnopharmacology*, 355(Pt B), 120704. <https://doi.org/10.1016/j.jep.2025.120704>.
- [10] Zhang, R., Zhu, X., Bai, H., & Ning, K. (2019). Network Pharmacology Databases for Traditional Chinese Medicine: Review and Assessment. *Frontiers in Pharmacology*, 10, 123. <https://doi.org/10.3389/fphar.2019.00123>.
- [11] Song, Z., Chen, G., & Chen, C. Y. (2024). AI empowering traditional Chinese medicine? *Chemical Science*, 15(41), 16844-16886. <https://doi.org/10.1039/d4sc04107k>.
- [12] He, Z., Liu, L., Han, D., Gao, K., Dong, L., Qi, X., Bu, D., Huo, P., Wang, Z., Fan, X., Sun, J., Lei, S., Deng, W., Liu, J., He, H., Zhao, Y., Guo, J. C., & Wu, Y. (2026). POINT: a web-based platform for pharmacological investigation enhanced by multi-omics networks and knowledge graphs. *Science Bulletin*, 71(3), 495-499. <https://doi.org/10.1016/j.scib.2025.10.002>.
- [13] Zhang, O., Lin, H., Zhang, X., Wang, X., Wu, Z., Ye, Q., Zhao, W., Wang, J., Ying, K., Kang, Y., Hsieh, C. Y., & Hou, T. (2025). Graph Neural Networks in Modern AI-Aided Drug

- Discovery. Chemical Reviews, 125(20), 10001-10103. <https://doi.org/10.1021/acs.chemrev.5c00461>.
- [14] Zhang, L., Shi, X., & Deng, X. (2025). MDL-HTI: A Multimodal Deep Learning Approach for Predicting Herb-Target Interactions. *Interdisciplinary Sciences: Computational Life Sciences*. <https://doi.org/10.1007/s12539-025-00772-w>.
 - [15] Wang, H., Xu, M., Guo, Z., You, G., Wang, G., Cao, C., & Hu, X. (2025). DTI-BGCGCN: A novel bipartite graph-based cluster graph convolutional network for drug-target interaction prediction in modern and traditional chinese medicine. *Computational Biology and Chemistry*, 117, 108410. <https://doi.org/10.1016/j.compbiolchem.2025.108410>.
 - [16] Hong, L., Teng, M., He, M., Zhao, J., & Li, S. (2026). AI-driven multi-scale target analysis of traditional Chinese medicine: From the pharmacological effects of single compounds to the synergistic mechanisms of formulae. *Journal of Advanced Research*. <https://doi.org/10.1016/j.jare.2026.03.050>.
 - [17] Wang, Y., Wu, T., Li, X., Xu, Q., Yu, H., Cen, S., Wang, Y., & Li, Z. (2025). Exploring artificial intelligence approaches for predicting synergistic effects of active compounds in traditional Chinese medicine based on molecular compatibility theory. *Chinese Journal of Natural Medicines*, 23(11), 1409-1424. [https://doi.org/10.1016/S1875-5364\(25\)60967-8](https://doi.org/10.1016/S1875-5364(25)60967-8).
 - [18] Niu, T., Zhu, Y., Mou, M., Fu, T., Yang, H., Sun, H., Liu, Y., Zhu, F., Zhang, Y., & Liu, Y. (2025). Identification of natural product-based drug combination (NPDC) using artificial intelligence. *Chinese Journal of Natural Medicines*, 23(11), 1377-1390. [https://doi.org/10.1016/S1875-5364\(25\)60942-3](https://doi.org/10.1016/S1875-5364(25)60942-3).
 - [19] Jin, S., Hong, Y., Zeng, L., Jiang, Y., Lin, Y., Wei, L., Yu, Z., Zeng, X., & Liu, X. (2023). A general hypergraph learning algorithm for drug multi-task predictions in micro-to-macro biomedical networks. *PLOS Computational Biology*, 19(11), e1011597. <https://doi.org/10.1371/journal.pcbi.1011597>.
 - [20] Wu, Y. P., Cheng, J., Sun, M., Qiao, C., Feng, Q., Fang, N., Ge, Y., & Rui, M. (2025). Artificial intelligence-driven identification and mechanistic exploration of synergistic anti-breast cancer compound combinations from *Prunella vulgaris* L.-*Taraxacum mongolicum* Hand.-Mazz. herb pair. *Frontiers in Pharmacology*, 15, 1522787. <https://doi.org/10.3389/fphar.2024.1522787>.
 - [21] Feng, C., Cheng, J., Sun, M., Qiao, C., Feng, Q., Fang, N., Ge, Y., & Rui, M. (2025). Artificial intelligence-driven identification and mechanistic exploration of synergistic anti-breast cancer compound combinations from *Prunella vulgaris* L.-*Taraxacum mongolicum* Hand.-Mazz. herb pair. *Frontiers in Pharmacology*, 15, 1522787. <https://doi.org/10.3389/fphar.2024.1522787>.
 - [22] Liu, J., Zhu, X., Zhang, X., Xu, Z., Wu, Z., & Jin, Z. (2025). AAFM-YOLO: TCM Identification Algorithm Based on Improved YOLOv11. 2025 5th International Conference on Electronic Information Engineering and Computer Science (EIECS), 264-267. <https://doi.org/10.1109/EIECS67708.2025.11283450>.
 - [23] Gao, S., Zhou, Z. Q., Zhong, S. Y., et al. (2025). Research on an Intelligent Identification Model for Traditional Chinese Medicine Decoction Pieces Based on Improved YOLOv3. *World Science and Technology-Modernization of Traditional Chinese Medicine*, 27(2), 364-374.
 - [24] Nie, W., Yuan, H., Ren, C. X., et al. (2025). Research on Quality Grade Classification of *Angelica dahurica* Decoction Pieces Based on Machine Learning. *Chinese Materia Medica and Clinical Practice*, 16(6), 1-6.
 - [25] Pan, Y., Wang, S., Ming, K., Liu, X., Yu, H., Du, Q., Deng, C., Chi, Q., Liu, X., Wang, C., & Xu, K. (2025). Leveraging AI technology for distinguishing *Eucommiae Cortex* processing levels and evaluating anti-fatigue potential. *Computers in Biology and Medicine*, 184, 109408. <https://doi.org/10.1016/j.compbiomed.2024.109408>.
 - [26] Nguyen Nhut, D. T., Tan, T. D., Quoc, T. N., & Hoang, V. T. (2024). Medicinal plant recognition based on Vision Transformer and BEiT. *Procedia Computer Science*, 234, 188-195.
 - [27] Li, X., Cheng, A. J., Liu, X. F., et al. (2025). Identification of *Cordyceps sinensis* authenticity based on convolutional neural network models. *Traditional Chinese Medicine and Clinical Practice*, 16(5), 8-12, 16.
 - [28] Gao, S., Zhang, H., Yang, Z., Wu, R., Chen, M., Zhai, D., Yang, Y., Qin, Y., Tao, H., & Li, P. (2026). A dual-branch CNN-encoder model for rapid geographic traceability of *Astragalus* using near-infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 346, 126851. <https://doi.org/10.1016/j.saa.2025.126851>.
 - [29] Shao, Y., Chen, D., Wang, B., Zhao, C., Tang, Y., Peng, Y., Zhang, H., Zhu, Y., & Tang, W. (2026). Identification of *Pueraria lobata* origin using terahertz precision spectroscopy and CNN-transformer hybrid network algorithm. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 348(Part 1), 127212. <https://doi.org/10.1016/j.saa.2025.127212>.
 - [30] Bai, R. B., Xiong, F., Wang, H., et al. (2025). Fusion of hyperspectral imaging color and spectral features based on deep learning for identification of *Salvia miltiorrhiza* cultivation areas. *Science of Traditional Chinese Medicine*, 3(3), 250-258.
 - [31] Yu, P., Yan, X., Hong, J., Li, D., Xiong, F., Bai, R., Yang, J., & Guo, L. (2025). Geographical origins of *Angelica sinensis* using functional compounds and multielement with machine learning-based fusion approaches. *Food Chemistry*, 471, 142747. <https://doi.org/10.1016/j.foodchem.2024.142747>.
 - [32] Nie, W., Huang, X. L., Pei, J., Ren, C. X., Zhou, T., Du, J., Jiang, H. J., Zhang, H. Y., Li, X., Li, J., Li, Y. H., Hu, Y. Y., & Hao, Z. Y. (2026). Research on quality evaluation system and grade classification of *Angelica dahurica* based on artificial intelligence and multispectral technology. *Chemometrics and Intelligent Laboratory Systems*, 269, 105610.
 - [33] Yang, J., Chen, K., Zhou, Y., Zheng, J., Sun, L., Zhang, Y., & Zhou, Z. (2025). Research on Geographical Origin Traceability of *Salvia miltiorrhiza* by Combining Two-Trace Two-Dimensional (2T2D) Correlation Spectroscopy and Improved DeiT Model. *Plants*, 14(21), 3365. <https://doi.org/10.3390/plants14213365>.
 - [34] Bai, R., Xiong, F., Wang, H., Luan, M., Zhou, J., Wan, X., Zhao, Z., Zhang, X., Zhang, C., & Yang, J. (2025). Deep learning-based fusion of color and spectral features from hyperspectral imaging for the origin identification of *Salvia miltiorrhiza*. *Science of Traditional Chinese Medicine*, 3(3), 250-258. <https://doi.org/10.1097/st9.0000000000000079>.
 - [35] Zhang, T. T., Tong, J. P., Li, L., Yuan, J. D., Yuan, X. T., Yuan, F. Y., Zhang, Z. Y., Gu, L., & Li, M. J. (2025). Integration of SWIR hyperspectral imaging, UHPLC-Q-TOF/MS and optimized machine learning for non-destructive authentication of *Pseudostellaria heterophylla*. *Industrial Crops and Products*, 234, 121534. <https://doi.org/10.1016/j.indcrop.2025.121534>.
 - [36] Liu, Y., Zhang, L. W., Zhang, X. Z., Bian, X. H., & Tian, W. L. (2025). Modern spectroscopic techniques combined with chemometrics for process quality control of traditional Chinese medicine: A review. *Microchemical Journal*, 213, 113605.
 - [37] Yang, B., Song, D., Li, Y., & Wang, J. (2025). Drug-drug interaction prediction of traditional Chinese medicine based on graph attention networks. *Scientific Reports*, 15(1), 18635. <https://doi.org/10.1038/s41598-025-00725-9>.

- [38] Ma, L., Su, C., Wang, J., Li, X., Sun, R., & Yu, G. (2025). AN-IHP: Incompatible Herb Pairs Prediction by Attention Networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 22(1), 124-135. <https://doi.org/10.1109/TCBBIO.2024.3498094>.
- [39] Zhang, R., Xie, W., Lin, Q., Li, X., & Wong, K. C. (2025). When East Meets West: Cross-Domain Drug Interaction Annotations With Large Language Models and Bidirectional Neural Networks. *IEEE Journal of Biomedical and Health Informatics*, 29(10), 7694-7703. <https://doi.org/10.1109/JBHI.2025.3563289>.
- [40] Li, Y., Zhang, X., Chen, Z., Yang, H., Liu, Y., Wang, H., Yan, T., Xiang, J., & Wang, B. (2024). Accurate prediction of drug-target interactions in Chinese and western medicine by the CWI-DTI model. *Scientific Reports*, 14(1), 25054. <https://doi.org/10.1038/s41598-024-76367-0>.
- [41] Lv, Q., Chen, G., He, H., Yang, Z., Zhao, L., Zhang, K., & Chen, C. Y. (2023). TCMBank-the largest TCM database provides deep learning-based Chinese-Western medicine exclusion prediction. *Signal Transduction and Targeted Therapy*, 8(1), 127. <https://doi.org/10.1038/s41392-023-01339-1>.
- [42] Weng, H., Xiao, B., Zhou, W., & Luo, Y. K. (2024). Real-World Drug-Adverse Reaction Prediction of Combined use of TCM Formula Granules and Western Medicine. 2024 11th International Conference on Behavioural and Social Computing (BESC), 1-9. <https://doi.org/10.1109/BESC64747.2024.10780482>.
- [43] Luo, Y. K., Wu, Q. B., Lai, W. H., et al. (2025). Research on an electronic medical record data mining method for Xuebijing Injection based on RAG-NLP. *China Journal of Health Information Management*, 22(6), 1007-1015.
- [44] Yang, H. Z. (2025). Large language models as a tool for molecular toxicity prediction: Insights from artificial intelligence into cardiac toxicity. *Journal of Chemical Information and Modeling*, 2268-2282. <https://doi.org/10.1021/acs.jcim.4c01371>.
- [45] Connor, S., Wu, L., Roberts, R. A., & Tong, W. (2025). Is ChatGPT Ready for Public Use in Organ-Specific Drug Toxicity Research? *Drug Discovery Today*, 30(2), 104297. <https://doi.org/10.1016/j.drudis.2025.104297>.
- [46] Li, Y., Liu, X., Zhou, J., Li, F., Wang, Y., & Liu, Q. (2025). Artificial intelligence in traditional Chinese medicine: advances in multi-metabolite multitarget interaction modeling. *Frontiers in Pharmacology*, 16, 1541509. <https://doi.org/10.3389/fphar.2025.1541509>.