Chinese Journal of Natural Medicines

Exploring artificial intelligence approaches for predicting synergistic effects of active compounds in traditional Chinese medicine based on molecular compatibility theory

Yiwen Wang, Tong Wu, Xingyu Li, Qilan Xu, Heshui Yu, Shixin Cen, Yi Wang, Zheng Li

Citation: Yiwen Wang, Tong Wu, Xingyu Li, Qilan Xu, Heshui Yu, Shixin Cen, Yi Wang, Zheng Li, 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, 2025, 23(11), 1409−1424. doi: 10.1016/S1875−5364(25)60967−8.

View online: https://doi.org/10.1016/S1875−5364(25)60967−8

Chinese Journal of Natural Medicines 23 (2025) 1409-1424

Contents lists available at ScienceDirect

Chinese Journal of Natural Medicines

journal homepage: www.cjnmcpu.com/

Review

Exploring artificial intelligence approaches for predicting synergistic effects of active compounds in traditional Chinese medicine based on molecular compatibility theory

Yiwen Wang[a,Δ], Tong Wu[a,Δ], Xingyu Li[a], Qilan Xu[a], Heshui Yu[a, b, c], Shixin Cen[a, b, c], Yi Wang[d,*], Zheng Li[a, b, c, e, f,*]

[a] College of Pharmaceutical Engineering of Traditional Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China
[b] State Key Laboratory of Component-Based Chinese Medicine, Tianjin 301617, China
[c] State Key Laboratory of Chinese Medicine Modernization, Tianjin 301617, China
[d] College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China
[e] Tianjin Key Laboratory of Intelligent and Green Pharmaceuticals for Traditional Chinese Medicine, Tianjin 301617, China
[f] Haihe Laboratory of Modern Chinese Medicine, Tianjin, 301617, China

ARTICLE INFO

Article history:
Received 23 January 2025
Revised 23 March 2025
Accepted 12 June 2025
Available online 20 November 2025

Keywords:
Molecular compatibility theory
Synergy prediction of TCM compounds
Molecular drugs combination prediction
Artificial intelligence

ABSTRACT

Due to its synergistic effects and reduced side effects, combination therapy has become an important strategy for treating complex diseases. In traditional Chinese medicine (TCM), the "monarch, minister, assistant, envoy" compatibilities theory provides a systematic framework for drug compatibility and has guided the formation of a large number of classic formulas. However, due to the complex compositions and diverse mechanisms of action of TCM, it is difficult to comprehensively reveal its potential synergistic patterns using traditional methods. Synergistic prediction based on molecular compatibility theory provides new ideas for identifying combinations of active compounds in TCM. Compared to resource-intensive traditional experimental methods, artificial intelligence possesses the ability to mine synergistic patterns from multi-omics and structural data, providing an efficient means for modeling and optimizing TCM combinations. This paper systematically reviews the application progress of AI in the synergistic prediction of TCM active compounds and explores the challenges and prospects of its application in modeling combination relationships, thereby contributing to the modernization of TCM theory and methodological innovation.

1. Introduction

Against the backdrop of the increasing number of complex diseases and the continuous evolution of drug resistance mechanisms, modern medicine is increasingly emphasizing multi-drug combination therapy strategies to enhance treatment efficacy and reduce side effects [1-4]. However, accurately predicting the synergistic effects of drug combinations remains a major challenge due to factors such as dose dependency, intricate drug-drug interactions, and poorly understood mechanisms of action [5, 6]. In response, researchers have increasingly adopted artificial intelligence (AI)-based computational methods, which have demonstrated significant progress [7-14]. This success is largely attributable to the ability of AI technologies to leverage large-scale multi-omics, structural, and phenotypic data to uncover potential synergistic relationships and facilitate the identification of highly effective drug combinations with diverse chemical spaces in complex disease contexts, including cancer, infectious diseases, and neurological disorders [15-23].

Traditional Chinese medicine (TCM) has long regarded the compatibility of herbal medicines as a core therapeutic principle [24]. Among various theories, the "monarch, minister, assistant, envoy" model provides empirical guidelines for the classification of drug functions, mutual coordination, and overall therapeutic harmony. TCM formulas (TCMFs), widely used in clinical practice [25], embody the synergistic mechanisms of multiple compounds and targets [26–27]. The potential pharmacological interaction networks within these formulas constitute a natural multi-drug synergy network [28]. However, due to the large number of active compounds and the complexity of their targets and signaling pathways, traditional empirical formulations struggle to elucidate their underlying molecular mechanisms [29–30]. As a result, applying modern technological approaches to model and validate the "monarch, minister, assistant, envoy"compatibility theory at structural and mechanistic levels has become a key challenge in the modernization of TCM. In response, the theory of molecular compatibility has emerged, aiming to transform traditional principles into molecular formulations with clearly defined structures and mechanisms [31-34]. This approach seeks to establish a unified framework linking compound structure, pharmacological function, and therapeutic mechanism in TCM [35-38], thereby

* Corresponding author.
E-mail addresses: zjuwangyi@zju.edu.cn (Y. Wang); lizheng@tjutcm.edu.cn (Z. Li)
Δ These authors contributed equally to this work.

https://doi.org/10.1016/S1875-5364(25)60967-8

providing a theoretical foundation for AI-based modeling.

However, the application of AI in the collaborative modeling of TCM remains in its infancy, facing significant challenges such as high data heterogeneity, limited model interpretability, and difficulties in quantifying traditional theories [39-40]. Building upon molecular compatibility theory, this article proposes a research framework that integrates TCM knowledge, molecular mechanisms, and AI technologies. It systematically organizes core concepts such as the "monarch, minister, assistant, envoy" hierarchy within the compatibility system and elucidates their relevance to modern drug synergy and molecular compatibility modeling. Furthermore, it examines the adaptability of mainstream AI algorithms in predicting the synergistic effects of active TCM compounds, and discusses strategies for managing heterogeneous data and enhancing model interpretability. Emphasis is placed on analyzing the potential and pathways for AI to simulate the intricate structure of the "monarch, minister, assistant, envoy" model. Through the integration of theoretical foundations, technological tools, and data resources, this study aims to advance the innovative development of TCM theory in the context of modern science and technology.

2. From classical compatibility to molecular modeling: AI-driven exploration of the "monarch, minister, assistant, envoy" theory

The theory of herbal formula in TCM is a concentrated expression of the TCM principle of syndrome differentiation and treatment. Its core lies in the synergistic effects of multiple herbs to achieve a holistic therapeutic effect [41-42]. The "monarch, minister, assistant, envoy" compatibility theory first appeared in the classical TCM text the *Yellow Emperor's Inner Canon* (Huangdi Neijing) [43]. This formula theory was refined through the clinical practice of subsequent generations of TCM practitioners, eventually forming a relatively complete system of "monarch, minister, assistant, envoy" compatibility theory [43-45]. The monarch medicine in a prescription is an indispensable core medicine that plays a major therapeutic role in the treatment of the main disease, and its medicinal power is the strongest in the prescription. The minister medicine is used to assist the monarch medicine in strengthening the treatment of the main disease or plays a major therapeutic role in important complications. Its medicinal power is weaker than that of the monarch medicine. The assistant medicine is auxiliary, an assistant medicine, or a counter-assistant medicine, and its medicinal power is weaker than that of the minister medicine. The envoy medicine is a guiding medicine or a harmonizing medicine, and its medicinal power is weak.

The "monarch, minister, assistant, envoy" compatibility theory of TCMFs clarifies the primary and secondary roles, division of labor, and coordination of drug combinations, allowing the formula to exert a comprehensive preventive and therapeutic effect [45]. In essence, this is not a simple arrangement of drugs, but a logical sorting of the condition, achieving overall intervention in the cause and location of the disease through the combination of drug properties [30]. This structured approach to drug use not only reflects the holistic treatment concept of TCM, but also provides a theoretical template for model analysis. The corresponding formula entities transform the idea of combination into specific drug combinations and dosages. Their design is based on a holistic understanding of the condition and demonstrates the synergistic effects of the active compounds [42]. Each formula contains the logic of mutual promotion and restraint between drugs, which is a direct reflection of the wisdom of TCM combinations.

Building on the hierarchical framework of drug functions represented by the "monarch, minister, assistant, and envoy" compatibility theory, modern pharmacological research seeks to elucidate the compatibility mechanisms of TCMFs at the molecular level [46]. This approach aims to tackle challenges such as the intrinsic complexity of TCM, the presence of multiple interacting compounds, the difficulty in identifying active compounds, and the lack of clear mechanistic interpretation within the framework of modern biomedicine [30, 47-48]. As shown in Fig. 1, the molecular compatibility theory was developed based on the traditional compatibility concept, aiming to clarify the active compounds in TCM and their mechanisms of action [35]. Through systematic identification and optimization of structure, efficacy, and targets, it achieves organic synergy among multiple compounds and constructs "molecular formulas" based on modern science [37].

Fig. 1 The transition from the "monarch, minister, assistant, envoy" compatibility theory to AI-driven molecular compatibility theory. The figure shows the process of transforming the traditional "monarch, minister, assistant, envoy" compatibility theory into molecular compatibility theory, which involves extracting active compounds from TCMs, combining them using AI based on the "monarch, minister, assistant, envoy" theory according to their specific structures, known effects, and mechanisms of action, determining the optimal dosage and efficacy, forming molecular formulations with good synergy and low toxicity, and then manufacturing modern TCMs using advanced pharmaceutical processes.

In recent years, modern TCM developed based on the molecular compatibility theory has achieved clinical success [36, 49]. Take Xie Tian's team's olibanum liposome as an example [32-33, 50-58]. *β*-olibanum is used as the monarch medicine due to its significant anti-tumor effect, while *γ* and *δ*-olibanum are used as minister medicines to synergistically enhance the efficacy of the medicine, and the envoy medicine is used to achieve precise delivery of the efficacy through the construction of a carrier. This division of roles at the molecular level not only demonstrates the extension of traditional combination thinking in the modern drug system, but also provides a basic template for structured modeling of AI models. Another classic case is Chen Zhu's team's compound Huang Dai Tablets, which achieved a complete remission rate of 96.7%-98% and a 5-year disease-free survival rate of 86.88% in the treatment of acute leukemia [36, 59-60]. The tetrasulfide in the compound is the monarch medicine, which directly degrades can-

cer proteins; the indigo red and tanshinone II A in the compound are minister medicines, which promote cancer protein degradation and cell differentiation, and inhibit the cell cycle and division and proliferation of cancer cells; Indigo also acts as an assistant medicine to reduce the toxicity of the monarch medicine, while tanshinone and indigo can also increase the number of channel proteins that transport arsenic tetrasulfide, increase the concentration of arsenic tetrasulfide entering leukemia cells, and play the role of envoy medicine.

These "monarch, minister, assistant, envoy" combinations at the molecular level forms a complementary network of medicinal effects and is a model for the "monarch, minister, assistant, envoy" relationship in modern Chinese medicine [34]. This method of combining active compounds in TCM according to the "monarch, minister, assistant, envoy" theory not only adheres to the holistic view of TCM and gives full play to the comprehensive therapeutic advantages of TCMFs' multiple targets, multiple links, multiple channels, and multiple active compounds, but also intuitively demonstrates the mechanism of action and rationality of combinations of active compounds in TCM at the cellular and molecular levels [61-65], overcoming the shortcomings of traditional TCMFs' unclear active compounds and unknown mechanisms of action. Most importantly, this theory transforms traditional empirical knowledge from vague, qualitative expressions into a modern scientific framework that is quantifiable and verifiable at the molecular level, thereby providing a structured and operational modeling theoretical basis for AI in Chinese medicine compatibility theory.

Currently, in the field of western medicine, research on drug synergistic effects has established a relatively mature modeling system, particularly in the prediction of binary synergistic combinations [10–11]. Researchers have developed a standardized evaluation system based on mathematical scoring models such as Bliss [66], Loewe [67], and HSA [6, 68]. These systems are primarily used to determine whether the combination of two drugs produces an effect exceeding the additive effect of using them separately. Furthermore, AI-based drug synergistic prediction models such as DeepSynergy [69], MatchMaker [70], and TranSynergy [71] have been developed. With the deepening of research, the focus of synergistic modeling has gradually expanded from binary combinations to ternary and higher-order multi-drug synergies, and ultimately to multi-drug combination therapy modeling used in real clinical practice [72]. In this context, the multi-compounds multi-level structure and functional characteristics of TCMFs provide a natural complex interactive system for AI to model drug synergistic relationships.

The concept of modeling synergistic interactions among active compounds in TCM draws inspiration from synergy research in modern pharmacology [27]. In western medicine, a relatively mature modeling system has been developed, progressing from binary combination scoring to high-order interaction modeling [73]. This reflects an ongoing evolution toward mechanism-oriented approaches. While TCMFs differ from western multi-drug regimens in origin and theory, they share key structural and functional parallels [27]. At the molecular level, TCMFs can be considered analogous to combination therapies, where multiple active compounds interact in ways comparable to complex multi-drug synergies [74]. The synergistic relationships among specific compounds in TCMFs also exhibit structural and functional similarities to binary drug combinations in modern medicine. As illustrated in Fig. 2, integrating a comprehensive understanding of the structural and multi-omics characteristics of active compounds in TCM with AI-based modeling frameworks from western medicine enables the systematic prediction of synergistic effects among active compounds in TCM. AI can identify potential synergistic interactions between these compounds, offering a foundation for the development of a functional drug hierarchy, role-based interaction network, and dose control model. This approach presents a viable pathway for computational modeling of the "monarch, minister, assistant, envoy"compatibility theory.

3. Basic data resources and evaluation standards for synergistic prediction of active compounds in TCM

The actual effectiveness of AI methods in predicting synergistic effects between active compounds in TCM depends not only on the architecture and learning ability of the model itself, but also heavily on the quality of the training data used and the scientific nature and suitability of the evaluation criteria [75–76]. TCM data exhibits distinct characteristics from western medicine data due to its extensive historical span, heterogeneous knowledge systems, and complex structure [77–78]. In this context, integrating classical literature, modern multi-omics data, and pharmacology research findings to establish a TCM data system that can be efficiently processed by AI models, and developing corresponding synergistic scoring frameworks constitutes foundational work for advancing TCM synergistic intelligent modeling.

3.1. Characteristics and challenges of TCM data

The task of modeling the synergistic effects between active compounds in TCM is significantly multi-level and multi-modal. Because its synergistic mechanism covers multiple levels, from molecular structure complementarity and target interaction to pathway regulation and phenotypic response, it relies on the comprehensive input of structural, functional, and semantic information [79]. As shown in Table 1, different types of data play different roles in modeling, from drug structure and composition, compound activity, to multi-omics data, TCM knowledge, and clinical phenotypes. Molecular structure data supports graph representation learning and feature extraction of the structure of active compounds in TCM, which is the underlying data foundation for AI modeling. Target and pathway information helps to construct a network of active compounds in TCM, targets, and diseases, supporting structural learning methods such as graph neural networks; while multi-omics data such as genomes, transcriptomes, proteomes, and metabolomes reflect the systemic response characteristics after TCM intervention and are suitable for modeling the overall effects of drug combinations in biological systems [79–80]. More importantly, TCM knowledge contains traditional semantic labels such as "monarch, minister, assistant, envoy" structure, "medicinal properties", and "contraindications". Although this type of knowledge is difficult to quantify directly, it can be converted into AI model inputs through knowledge representation and role embedding. This knowledge not only carries the theoretical framework of TCM, but also provides the model with semantic-oriented "structural guidance", which is particularly important for constructing explanatory models with TCM logic [81].

Unlike the structured, standardized, and clearly labeled data systems in western medicine research, TCM data is characterized by diverse sources, complex structures, and semantic ambiguity, posing numerous challenges for modeling. For example, classical texts contain extensive natural language descriptions that lack unified terminology control and standardized structures [76]. Modern TCM research offers target and pathway information but suffers from data dispersion and inconsistent naming conventions [29]. In addition, multi-omics and clinical data are often not systematically archived, making them difficult to directly use for model training [92–93].

In response to the challenges, as shown in Table 2, several TCM databases have attempted to achieve structured integration of basic information. Platforms such as TCMSP [94], BATMAN-

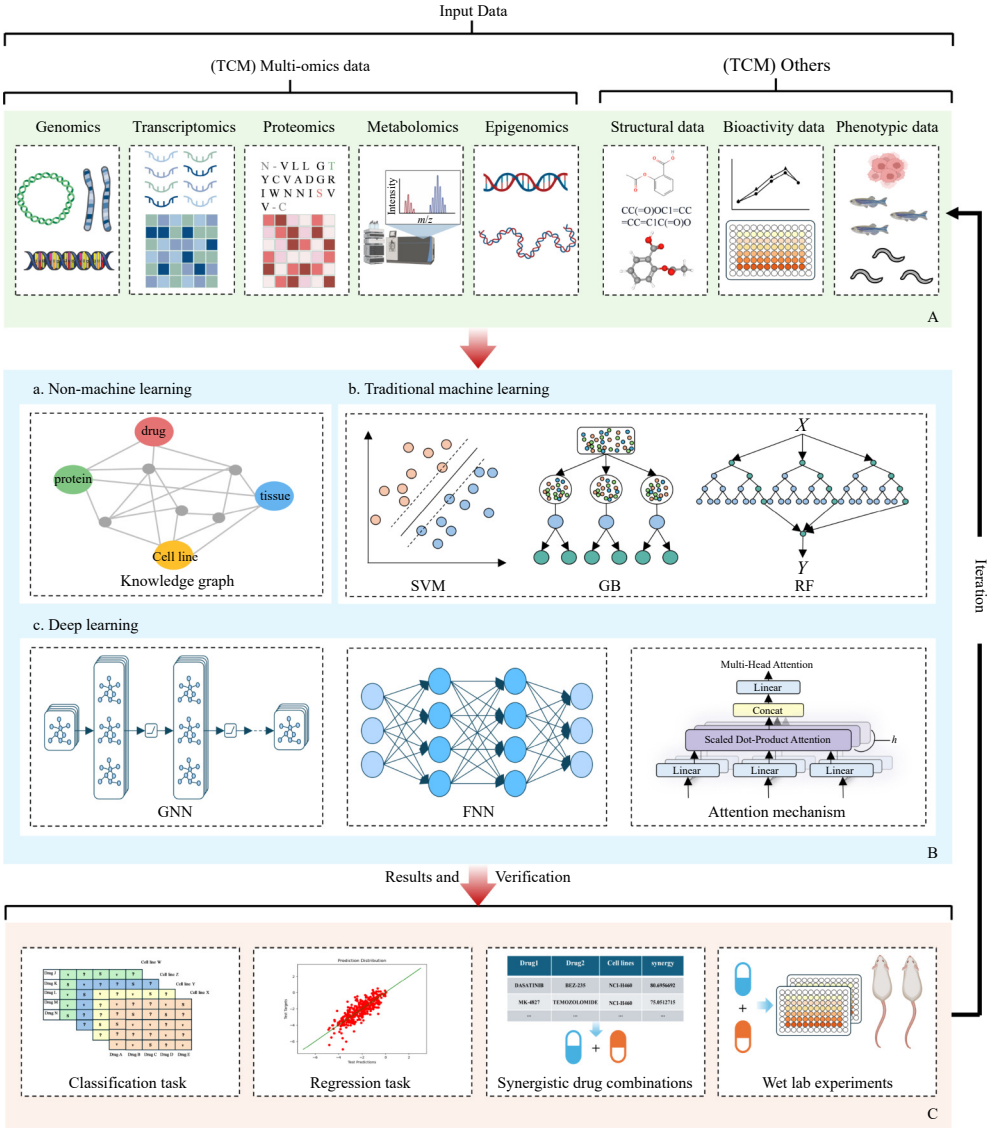

Fig. 2 Flow chart for predicting synergistic effects of drug combinations at the molecular level. This figure presents an integrated workflow for drug synergy prediction (in TCM), covering the entire process from data acquisition to closed-loop optimization. (A) It begins by collecting and preprocessing input data (in TCM), including molecular profiles, multi-omics, drug combination experiments, and phenotypic responses. (B) A core predictive model is constructed using either traditional ML methods or DL approaches. (C) This model is trained on labeled data and evaluated through cross-validation and ROC-AUC metrics. The selected combinations undergo external validation and are further confirmed by wet-lab experiments both *in vitro* and *in vivo*, as well as through clinical trials, with parallel exploration of underlying biological mechanisms. Experimental outcomes are feedback to refine model parameters, completing the loop to enhance prediction accuracy and accelerate the development of effective drug combinations.

TCM [95], ETCM [96], and HERB [97] cover topics such as herbal medicine-compound-target, pharmacologic activity prediction, drug meridian affiliation, and formula structure, respectively. Some databases have also introduced network pharmacology association maps. Additionally, platforms like TCM-Mesh [98] integrate omics data with target networks, providing foundational data support for studying synergistic mechanisms at the systems level. However, these databases generally suffer from issues such as delayed updates, incomplete coverage, and inconsistent data quality, particularly lacking specialized structured resources for synergistic labeling of TCM compound combinations. At the same time, with the widespread availability of multi-omics data, genomic, proteomic, metabolomic, and transcriptomic data are gradually being applied to the analysis of TCM synergistic mechanisms, providing biological support at the mechanism level for AI models. The integration of multi-omics and AI is expected to promote the transformation of TCMFs from empirical expression to mechanism modeling [99].

In summary, current TCM data exhibits typical characteristics of diverse sources and disorganized structures, as well as rich content and dispersed semantics, presenting challenges such as inconsistent input representations, insufficient samples, and scarce labels for AI modeling. This makes data preparation, representation, and conversion the primary tasks in TCM synergistic prediction modeling.

3.2. Modeling strategies and data organization methods

To meet the requirements of AI models for structured input and semantic consistency, researchers have proposed various modeling and organization strategies for TCM data to support effective prediction of synergistic effects. First, at the semantic level, unstructured text information must be converted into vector representations. Some pre-trained models, such as TCM-GPT [124], TCMChat [125], Zhongjing [126], trained on TCM corpora, can capture the semantic variability of drug terms in context, enabling contextual embedding representations of elements such as drugs, effects, and meridian tropism [127]. Additionally, using embedding methods like Node2Vec to establish semantic similarity measures between drugs and compounds can also introduce lat-

Table 1 Foundational data types supporting AI-based modeling of synergistic drug effects in TCM.

The types of data	Characteristics	Representation
Genomics [82]	Reflects the genetic background of cell sensitivity to drugs.	Mutations, copy number variation, single nucleotide polymorphism, etc.
Epigenomics [83]	Reveals the drug influence at the gene expression regulation level.	DNA methylation, histone modification, etc.
Transcriptomics [84]	Reflects the gene expression changes in cells before and after drug treatment.	Gene expression profiles, etc.
Proteomics [85–86]	An important mediator of drug-target interactions.	Protein expression and interaction networks.
Metabolomics [87]	Reflects the cell energy metabolism pathways and toxicity changes, suitable for multi-targets mechanism research.	Abundance of small-molecule metabolites.
Drug structure [88]	Describes the basic characteristics such as the molecular composition and structural features of drugs.	the simplified molecular input line entry system (SMILES), molecular fingerprints, physicochemical properties of drugs, etc.
Drug activity [89]	Directly reflects the actual biological effects when drugs are used alone or in combination.	Growth inhibition rate, IC50, synergy score, etc.
Phenotypic data [90]	Reflects the results of the interaction between genes and the environment.	Cell morphology, growth inhibition, mortality rate, metabolic changes, mouse behavior experiments, observation of abnormal zebrafish development, etc.
TCM knowledge [91]	Basic theories of TCM, principles of combination, and diagnostic and therapeutic methods.	Yin-yang and five elements theory, viscera-meridian theory, four natures and five flavors of herbs, "monarch, minister, assistant, envoy" formula principle, four diagnostic methods, acupuncture and moxibustion therapy, dietary therapy, etc.

ent associations between drugs in structural learning [128].

Furthermore, considering that TCM formulas inherently possess a multi-level structure, constructing a multi-level heterogeneous graph with nodes such as “compounds-target-disease” and edges such as “contains”, “acts on”, and “treats” has become the mainstream modeling approach [129-131]. This graph not only preserves the hierarchical and functional relationships within TCMFs, but also facilitates subsequent collaborative modeling and feature propagation of the graph structure using GNN-type methods, thereby revealing potential synergistic effect patterns [129].

It is worth noting that classic formulas themselves contain a wealth of synergistic experience information, and their drug combination patterns, combination ratios, and “monarch, minister, assistant, envoy” structures can provide valuable labeling systems for models. Through natural language processing, relationship extraction, or knowledge graph construction methods, typical synergistic drug pairs, functional role labels (such as monarch medicine/assistant medicine, etc.), dosage ratios, and four energies and five flavors can be extracted from the literature. For example, the AMFormulaS extracts formula compatibility information from ancient books, including drug names, dosages, and usage methods [132]. This process not only enhances the degree of data structuring, but also provides semantic support for the subsequent introduction of modeling elements such as role constraints and dosage conditions.

In summary, the restructuring of data should achieve three objectives: first, converting the semantic structure of TCM knowledge into vectors or graph structures recognizable by the model; second, integrating multi-source information to form upstream and downstream association graphs; and third, extracting implicit knowledge related to modeling from the classical TCM system to incorporate “TCM logic” into the model.

3.3. Drug synergy scoring criteria

Predicting synergistic effects of active compounds in TCM requires not only structured data support, but also scientific and reasonable synergy scoring criteria to evaluate and guide training based on the prediction results. In modern drug research, models such as bliss independence [66], loewe additivity [133], highest single agent [133], and zero interaction potency [134] are widely used to quantify the difference between the combined effect of two drugs and the effect of a single drug. These models are based on independent probability theory, dose addition hypothesis, or maximum single drug effect criteria, respectively, and can objectively measure whether there is a synergistic effect in the combination.

However, most of the above-mentioned scoring systems rely on accurate dose-response curves and control experiment data, which are the weakest links in the field of TCM [42]. Molecular formulas composed of active compounds in TCM are difficult to standardize in experiments due to their complex composition, dose variations, and diverse mechanisms. Therefore, the application of scoring models used to quantify synergistic effects in western medicine drug combinations faces significant limitations in TCM [27].

Currently, in TCM synergistic evaluation research, the following strategies are primarily employed to address this challenge [42]. First, knowledge-based indirect scoring, which projects TCM compounds onto existing western medicine target points or pathway systems and uses western medicine synergistic scoring results for comparison. Second, semantic scoring based on expert experience, which combines descriptions of certain formula effects in TCM literature to establish an expert rule database or scoring template. Third, experimental indicator construction based on multi-omics responses to quantify pharmacological synergy. Fourth, exploring unique scoring indicator systems applicable to the TCM system, such as integrating drug properties, dosage ratios, and formula structures to construct feature-driven synergy indices.

Although there is currently no unified scoring system for TCM synergistic effects, with the accumulation of data, the standardization of experimental protocols, and the increasing involvement of AI, establishing a mechanism-based evaluation system for TCM synergistic effects that integrates traditional knowledge systems has become a promising direction.

4. Analysis of the applicability of AI algorithms in the synergistic prediction of active compounds in TCM

4.1. Mainstream AI model principles and cases

4.1.1. Traditional machine learning algorithms

Traditional machine learning methods, including Random Forest (RF), Support Vector Machine [135] (SVM), and Gradient Boosting (GB), are characterized by relatively simple architectures and lower dependence on large-scale training data. The underlying principles of these algorithms are illustrated in Fig. 3, highlighting their core mechanisms. They demonstrate strong discriminative abilities in feature space and are particularly effective for drug synergy prediction under conditions of limited samples and low-dimensional inputs. For instance, SVM has been used to integrate multiple drug feature types to achieve reliable synergy prediction [136], while GB-based models such as PDC-SGB enhance predictive accuracy by integrating compound features,

target information, and pharmacological pathways [137]. Sun et al. employed RF to screen combinations of Chinese and Western compounds, identifying synergistic combinations such as myricetin with celecoxib and rhein with hydroxychloroquine, which exhibited strong synergistic effects (CI < 0.7) in RAW264.7 cells [138].

Despite these advantages, traditional machine learning methods struggle to model the nonlinear and hierarchical interactions characteristic of TCMFs. They lack the ability to capture structured semantic relationships such as the hierarchical roles of monarch, minister, assistant, and envoy, and their interpretability is limited without incorporating external prior knowledge. Consequently, while suitable for preliminary screening and small-scale analysis, these methods face challenges in addressing the complexity of TCM compatibility modeling [138].

4.1.2. Deep learning algorithms

Deep Learning is a machine learning method based on neural

Table 2 The summary of public databases containing the effective compounds of TCM.

Database Type	Database	Core Data	Unique Features	Website for database	Interoperability	References
TCM/Natural Product-Related Databases	TCMID	TCM compounds, targets, diseases, prescriptions, and related literature, etc.	Provides information and bridges the gap between TCM and modern life sciences.	http://www.megabionet.org/tcmid/	Supports linkage with databases such as UniProt and KEGG, allows export of CSV/TSV format data.	100
	TCMbank	Includes chemical structures, pharmacological effects, and clinical application data of TCM formulas, single herbs, and active compounds.	TCMBank provides literature-based herb/ingredient-gene/disease mappings to explore mechanisms and identify treatments, with its intelligent module continuously integrating PubChem-derived TCM data.	https://tcmbank.cn	Supports structure similarity search, interoperable with PubChem.	101
	SymMap	Integrates TCM compound, target, the clinical symptoms and disease to construct multi-scale biological networks.	Provides network visualization analysis, drug repositioning, and target prioritization tools	http://www.symmap.org	Provides network-based visualizations and user-defined GO/pathway enrichment analysis.	102
	HIT	Focuses on association data between herbal active compounds and targets, covering compound structures, target functions, and interaction information.	HIT uniquely supports automated target mining from daily PubMed updates and enables personalized online target curation through its My-target system.	http://hit2.badd-cao.net	Crosslinks were made to databases of TTD, DrugBank, KEGG, PDB, UniProt, Pfam, NCBI, TCM-ID and others.	103
	HERB	Contains experimental validation data for herbal active compounds, action targets, and disease associations.	Integrates high-throughput transcriptomic data with manually curated references.	http://herb.ac.cn/	Enhances interoperability by integrating and cross-referencing data from multiple TCM and biomedical databases, including HIT, TCMSP, SymMap, DrugBank, DisGeNet, and CMap.	97
	ETCM	Integrates TCM effective compounds, action targets, and pharmacokinetic parameters.	Provides ADME/Tox property prediction, target-disease association analysis, and formula compatibility analysis tools.	http://www.tcmip.cn/ETCM2	Supports data interaction with Pubchem, Uniprot and Genecards,	96
	TCMSP	Contains TCM compound, target, pathway, and pharmacokinetic data, constructing a TCM systems pharmacology network.	Provides compound screening (based on OB, DL, etc.), target prediction, and pathway enrichment analysis functions.	https://www.tcmsp-e.com/	The obtained network can be downloaded for further analysis.	94
Chemical Molecules and Drug Databases	DrugBank	Integrates drug chemical structures, target proteins, action mechanisms, and clinical information.	Provides drug-target interaction query, metabolic pathway analysis, and drug similarity search functions.	https://go.drugbank.com	DrugBANK's clinical API includes a robust drug interaction checker and advanced drug search features.	104
	ZINC	Contains small-molecule compound structures for virtual screening, supports fragment and combinatorial library searches.	Provides structure-based virtual screening tools, supports 3D structure generation and molecular docking parameter optimization.	http://zinc.docking.org/	Compatible with software such as PyMOL and AutoDock, allows download of SDF format files.	105
	PubChem	Covers compound structures, properties, biological activities, and related literature data.	Provides multi-dimensional search functions for structure, properties, and biological activity, as well as cheminformatics tools and data visualization functions.	https://pubchem.ncbi.nlm.nih.gov/	Supports integration with NCBI Gene and PubMed, allows export of multiple formats.	106
Compound-Target Association	TTD	Contains validated and potential therapeutic targets, covering target-disease associations and drug development status.	Provides target screening (based on disease type, drug status), target function annotation, and drug development pipeline analysis	https://db.idrblab.net/ttd/	TTD provides API, allowing users to download data in batches and conduct customized analysis.	81
	STITCH	Integrates protein-compound interaction data, covering direct binding and functional associations.	Provides interaction network visualization, functional enrichment analysis, and compound-target association prediction tools.	http://stitch.embl.de	Supports integration with STRING protein interaction networks, allows export of network data.	107
	BindingDB	Stores small molecule-protein binding affinity data, including KD, IC_{50}, and other parameters.	Provides binding data retrieval, affinity ranking, and molecular docking result validation functions.	https://www.bindingdb.org	Supports interoperability with Reactome, provide data download and API.	108

Continued

Database Type	Database	Core Data	Unique Features	Website for database	Interoperability	References
Protein-Protein Interaction and Related Information Databases	UniProt	Integrates protein sequence, functional annotation, post-translational modification, and interaction data.	Provides protein function query, sequence alignment, and annotation analysis tools, supports customized data download.	https://www.uniprot.org/	Supports association with PubMed and PDB, allows export of FASTA/XML formats.	109
	STRING	Predicted and experimentally validated protein interaction networks, covering multi-species data.	Provides interaction network visualization, functional module identification, and enrichment analysis tools, supports cross-species comparison.	https://cn.string-db.org/	Supports integration with KEGG and GO, allows export of networks for Cytoscape.	110
	MINT	Stores experimentally validated protein-protein interaction data, covering multiple species.	Provides interaction data retrieval, network analysis, and literature mining functions, supports data submission.	https://mint.bio.uniroma2.it/	Provide data download and API.	111
	IntAct	Integrates public protein interaction data, supports data validation and annotation.	Provides interaction data query, network visualization, and analysis tools, supports community-contributed data.	https://www.ebi.ac.uk/intact/home	Associated with UniProt and ChEMBL, allows export of multiple formats.	112
	BioGRID	Contains protein-protein and gene-gene interaction data for multiple organisms.	Provides interaction data retrieval, network analysis, and functional enrichment tools, supports batch data download.	https://thebiogrid.org/	Allows export of CSV/TSV formats.	113
Disease Target Databases	DisGeNET	Integrates disease-gene/protein association data, sourced from literature mining and genomic studies.	Provides disease-gene association retrieval, enrichment analysis, and network visualization tools.	https://www.disgenet.org/	Supports interoperability with Cytoscape, allows export of association data.	114
	CTD	Stores chemical-gene-disease association data, covering toxicology and environmental health information.	Provides chemical-gene-disease network analysis, pathway enrichment, and literature mining functions.	https://ctdbase.org/	Supports integration with GO, allows export of association tables.	115
Omics data	GDSC	Genomics of Drug Sensitivity in Cancer database, containing cell line drug response and genomic data.	Provides drug sensitivity query, genomic variation and drug response association analysis tools.	http://www.cancerrxgene.org/	Allows export of drug sensitivity data.	116
	CCLE	Cancer Cell Line Encyclopedia, covering gene expression, mutation, and copy number variation data of cell lines.	Provides cell line characteristic retrieval, gene expression analysis, and drug sensitivity prediction tools.	http://www.broadinstitute.org/ccle	Provide data download and API.	117
	COSMIC	Catalogue of Somatic Mutations in Cancer, containing gene mutation data in human cancers	Provides mutation retrieval, tumor type distribution analysis, and mutation function annotation tools	http://www.sanger.ac.uk/cosmic	Allows download of mutation tables.	118
	ICGC	International Cancer Genome Consortium database, integrating multi-omics cancer data (genomics, transcriptomics, etc.).	Provides cancer multi-omics data retrieval, visualization, and integrated analysis tools, supports data download.	https://platform.icgc-argo.org/	Supports CSV/TSV format output.	119
	TCGA	Covers multi-omics data (clinical, genomic, expression, epigenetic, etc.) of 33 cancer types, one of the most authoritative data sources in cancer research.	Provides multi-omics data retrieval, integrated analysis, and visualization tools, supports data mining.	https://portal.gdc.cancer.gov/	Supports API interface and standardized data formats.	120
	Harmonizome	Provides nearly 12 million gene associations with various attribute types (such as cells and tissues, diseases, and pathways).	Harmonizome 3.0 also supports cross-dataset analysis, LLM-assisted interpretation, diverse visualization, and provides rich omics data in AI-ready formats.	https://maayanlab.cloud/Harmonizome/	Supports interaction with various omics databases, allows export of AI-compatible format data.	121
	MLOmics	Integrates 4 types of omics data covering 32 cancer types and 8,314 samples.	Provides 20 classification, clustering, and imputation tasks, statistical/ML/DL baselines, links to biological knowledge databases, and supports downstream toolchains such as survival analysis.	https://github.com/chenzRG/Cancer-Multi-Omics-Benchmark	Supports compatibility with common machine learning frameworks, allows download of standardized datasets.	122
	scImmOmics	Integrates data from 7 single-cell sequencing technologies including scRNA-seq and scTCR-seq.	Provides multi-dimensional retrieval, UMAP visualization, multi-omics integration analysis (including transcriptome/epigenome/receptorome), and functional tools (cell communication/differentiation trajectory) to facilitate basic and clinical immune research.	https://bio.liclab.net/scImmOmics/home	Allows export of single-cell data.	123

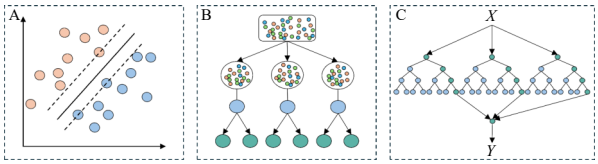

Fig. 3 Schematic diagram of classical Machine Learning algorithms. (A) The network architecture of support vector machine. (B) The network architecture of gradient boosting. (C) The network architecture of random forest.

networks, which excels at handling high-dimensional and unstructured data [139–140]. In the prediction of drug combination synergy, drug molecular features and cell line features are typically high-dimensional multimodal data. Traditional machine learning methods often struggle to fully extract the underlying information from these data, whereas deep learning, through multi-layer feature transformations and nonlinear representations, is capable of capturing the complex relationships between drugs and cell lines [141].

4.1.2.1 Deep belief networks

As shown in Fig. 4A, deep belief networks (DBNs) consist of stacked restricted Boltzmann machines and extract high-dimensional features through layer-wise pretraining [142]. The DBN model by Chen et al. integrates gene expression, signaling pathways, and biological features to predict drug synergy with favorable outcomes [143]. DBNs possess strong generative and abstract feature learning capabilities, making them suitable for ambiguous or unlabeled TCM data, including classical texts and omics. However, they are prone to overfitting on small datasets, require complex parameter tuning, and lack interpretability and modeling capacity for TCM compatibility logic, limiting their application in TCM compound modeling.

4.1.2.2 Feedforward neural network

Feedforward neural networks (FNNs) construct a unidirectional computational path from input to output through multiple layers of nonlinear mappings, as shown in Fig. 4B, enabling them to capture complex relationships between high-dimensional features [144]. The DeepSynergy model developed by Kristina Preuer et al. predicts the synergistic score of anticancer drug combinations by inputting drug structural features and cell line expression profiles [69]. FNNs offer advantages such as structural simplicity and flexible training in TCM modeling. They can integrate structural features of active compounds with gene expression or multi-omics data from cell lines, in a manner similar to models like DeepSynergy and AuDNNsynergy [141] used for predicting drug combination synergy. Its limitations lie in its lack of structural modeling capabilities, making it difficult to express the hierarchical compatibility relationships in TCMFs, and its inability to reflect the "monarch, minister, assistant, envoy"functional dependencies. Furthermore, the model is sensitive to feature quality and parameters, and its explanatory power is weak. Overall, FNN serves as an effective tool for multimodal feature fusion. However, it lacks the ability to capture the hierarchical logic embedded in the structure of compound, role, and mechanism.

4.1.2.3 Encoder-decoder network

The encoder-decoder network achieves complex nonlinear

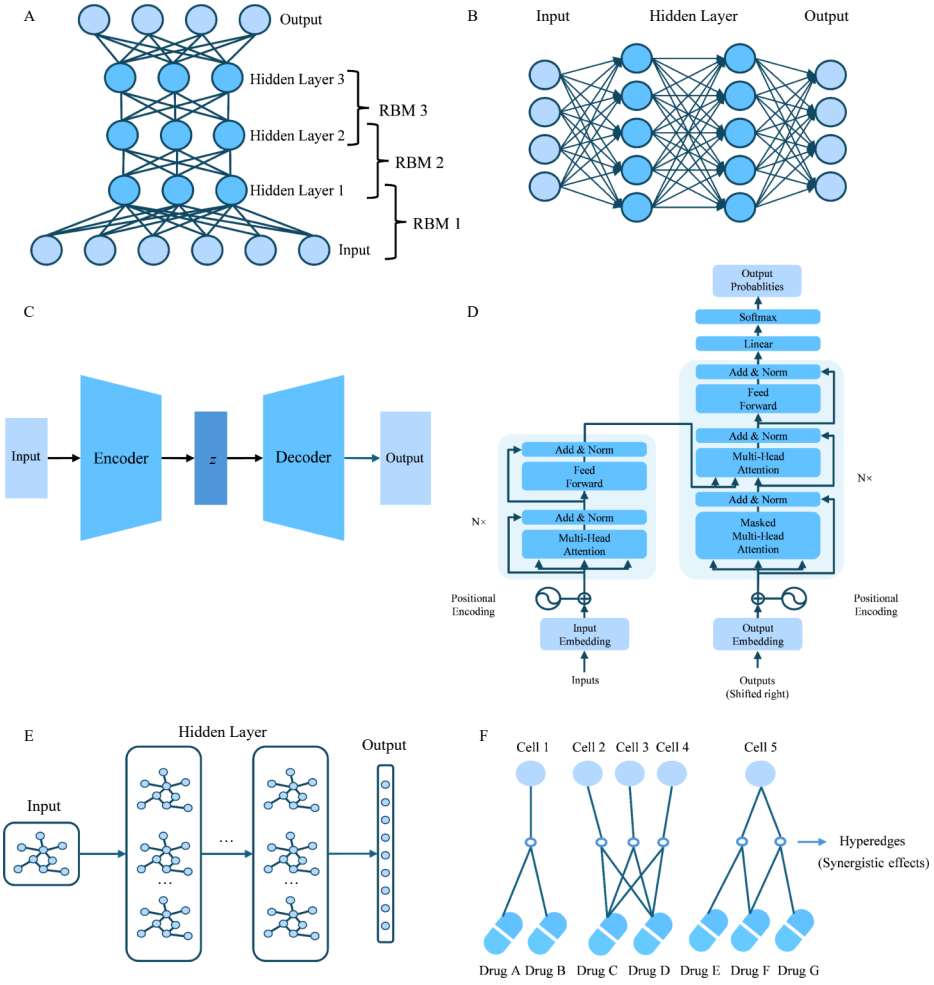

Fig. 4 Diagram of the principles of deep learning algorithms. (A) The architecture of deep belief network. (B) The architecture of feedforward neural network. (C) The architecture of encoder-decoder network. (D) The network architecture of transformer. (E) The architecture of graph neural network. (F) The relationship between drugs and cell lines based on hypergraph representation.

alignment between input and output by embedding input features into a high-dimensional latent space and then remapping them to the output [145] (Fig. 4C). SDCNet adopts a multi-channel encoder-decoder architecture, where graph-based encoders extract semantic representations from DDI, DTI, and DPI subgraphs, and a Transformer-based decoder, guided by cell line context, infers drug-drug synergy scores through attention-driven interaction modeling [146]. The advantage of this type of model in TCM modeling is its flexible structure, which facilitates the introduction of heterogeneous information such as "drug properties, meridian tropism, and role labels". At the same time, it supports collaborative classification, regression, and generation tasks and can be used to learn potential compatibility relationships from data. Its limitations are its complex model structure, high sensitivity to the number of training samples and feature coupling, and lack of structural expression paths for knowledge constraints in TCMFs, such as compounds priority and dose dependence.

4.1.2.4 Transformer-based models

Transformer models utilize self-attention mechanisms to model long-range dependencies, achieving breakthrough results in natural language processing before being gradually adopted for drug interaction prediction tasks [147-149]. They are particularly well-suited for handling high-dimensional, unstructured, and cross-modal data features [150]. Based on model scale and task complexity, transformer-based methods can be broadly categorized into non-LLM transformer models and large language models (LLMs), each exhibiting distinct characteristics in structural modeling and semantic fusion.

Non-LLM transformer models typically employ standard multi-head attention mechanisms and encoder structures to embed and model inputs such as drug structures. These architectures enable task-oriented feature extraction and relationship construction (Fig. 4D). Representative models include TransSynergy, which constructs gene-gene interaction networks and computes attention weights to uncover associations between drug synergies and signaling pathways [71]. SMILESynergy further incorporates SMILES strings as drug sequence inputs to enhance structure-dependent expression and modeling capabilities [88]. In the context of TCM modeling, these models offer the advantage of flexibly integrating non-structural semantic features such as drug properties, chemical constituents, and meridian associations. They support multi-source data fusion through multi-channel attention mechanisms and are well suited for capturing complex relationships among drugs, their pharmacological functions, and compound interactions. However, non-LLM models often lack a general understanding of cross-drug systemic semantics and provide limited support for knowledge abstraction or language-driven tasks in traditional texts.

Large language models (LLMs) extend the transformer structure to multimodal modeling based on large-scale pre-training and contextual learning capabilities and possess powerful semantic understanding and knowledge transfer capabilities. SynerGPT encodes drug combinations, cell lines, and synergistic scores as language sequences, learns synergistic rules through pre-training, and achieves "zero-shot prediction" for unknown combinations [151]. BAITSAO introduces a Shapley value interpretation mechanism on this basis, embedding semantic relationships between drugs, genes, and cells to complete multi-task modeling and synergistic mechanism interpretation [152]. LLMs have multiple advantages in TCM modeling: First, they can use medical texts, classic formulas, classic TCM works such as "compendium of materia medica", and other corpora for cross-contextual knowledge mining. Second, they can integrate text elements such as "drug properties, roles, dosages, and symptoms"to automatically generate structured embeddings for modeling "monarch, minister, assistant, envoy" style combination logic. Third, they possess the ability to integrate "language-graph-omics" multimodal inputs, making them suitable for constructing compound models with reasoning capabilities. However, their limitations are also evident: on the one hand, they have high training costs and significant computational power requirements and lack a dedicated TCM corpus system to support them; on the other hand, their explanatory power at the pharmacological mechanism level remains insufficient, making it difficult to accurately map the causal relationships between drug synergistic effects and biological pathways.

Overall, the transformer architecture has significant potential in TCM synergistic modeling, especially in modeling complex hierarchical relationships and text-structure-omics integration. In the future, if we can combine TCM knowledge graphs and semantic role ontologies to develop lightweight, specialized TCM language models, transformer is expected to become a key supporting framework for TCM compound AI modeling.

4.1.2.5 Graph neural networks

Graph neural networks (GNNs) are a class of deep learning algorithms specifically designed for processing graph-structured data, widely applied in scenarios such as biomolecular network modeling and drug combination prediction [153]. Their core principle involves aggregating and updating node features through graph convolution operations to learn underlying interaction patterns within network structures, particularly suited for capturing complex dependencies among biological entities such as drugs, compounds, and targets (Fig. 4E). Gu et al. developed a GNN model to predict TCM compatibility for colorectal adenoma treatment [154]. The study constructed a heterogeneous graph network involving herbs, small molecules, and targets, employing GCN and multi-layer perceptions for collaborative learning high-potential herb pairs screened by the herbal combination potential index were validated *via* molecular docking and CCK-8 cell proliferation assays, showing significant tumor cell inhibition. This research indicates that GNNs have inherent advantages in processing hierarchical and structured TCM data, particularly suitable for modeling herb-compounds-target ternary networks.

In drug synergy prediction tasks, typical GNN architectures include GCN [155-156], GAT [157], GAE, GraphSAGE, and HGNN. GCN extracts the topological structure and chemical features of drug molecular graphs through graph convolution, successfully achieving joint modeling of drug combinations and cellular responses in the DeepDDS [158]. GCN is suitable for molecular modeling tasks with clear TCM compound structures and can characterize the interactions between compounds in TCMs, but it has limited ability to express the hierarchical structure and high-order compatibility logic of the "monarch, minister, assistant, envoy" model, making it difficult to distinguish the role weights and regulatory pathways between compounds.

GAT dynamically allocates the weights of adjacent nodes by introducing an attention mechanism, strengthening the modeling ability of key compounds and their targets. The JointSyn model is based on GAT and multi-perspective learning strategies, which improve the generalization ability of unknown drug combinations [159]. GAT is suitable for expressing semantic logic such as "main drug orientation and assistant medicine toxicity" in combination structures, but due to the large computational overhead, training efficiency is limited by the complex attention mechanism, and the interpretability of the model needs to be improved.

GAE, as an unsupervised graph autoencoder structure, combines GCN or GAT encoders with structural reconstruction decoders and performs well in scenarios lacking labeled data. The GAECDS model utilizes GAE to learn Chinese medicine compound-target graph embeddings and achieves collaborative scoring prediction [160]. GAE is well suited to TCM research where labels are scarce and is suitable for modeling traditional literature and formula knowledge extraction, but its semantic logic modeling abil-

ity is weak and needs to be integrated with prior knowledge to enhance its explanatory ability.

GraphSAGE uses an inductive learning framework and neighborhood sampling mechanism to solve the scalability problem of GCN on large-scale graphs and has the ability to process graphs with millions of nodes [161]. In drug synergy tasks, GraphSAGE can quickly adapt to new compound nodes and support dynamic updating and iterative modeling of TCM knowledge graphs. It has high modeling efficiency and strong generalization capabilities, making it suitable for handling structural evolution problems in TCMFs, but it is still insufficient for expressing multi-level semantic relationships and modeling combination logic.

HGNNs are a type of deep learning method used to model complex, multi-dimensional relationships [162]. As shown in Fig. 4F, the core advantage of hypergraphs lies in their ability to represent multi-dimensional relationships. Their core lies in simultaneously connecting multiple nodes through hyperedges, thereby representing higher-order dependencies between multiple drug compounds, cell lines, or genes. Compared to traditional graph structures where each edge connects only two nodes, HGNNs have a natural advantage in capturing multi-compound interactions. The HypergraphSynergy model developed by Liu et al. represents the drug-drug-cell triplet as a hyperedge to predict synergistic combination effects, maintaining good performance even in the absence of combinations [162]. HGNN is particularly suitable for modeling complex, non-additive, multi-target synergistic mechanisms in TCMFs, and can be embedded with "monarch, minister, assistant, envoy" role information to form a network map with a Chinese medicine semantic structure. However, HGNN performance is limited when dealing with sparse data or lacking explicit role annotations. To address this issue, MHCLSyn introduces a multi-perspective contrastive learning mechanism and feature compression enhancement module, improving representation capabilities and prediction accuracy in sparse graphs [163]. The HGNN method is particularly suitable for characterizing the complex interactions in Chinese medicine molecular combinations from multiple levels of "structure-semantics-mechanism" and can provide a high-order topological expression framework for compound combination modeling. However, further exploration is still needed in the construction of a semantics-driven labeling system and the fusion of multi-source heterogeneous data.

Overall, GNNs provide powerful structured expression tools for TCM molecular compatibility modeling, and are particularly suitable for integrating multi-dimensional heterogeneous information such as compounds, targets, pathways, and roles to express the complex synergistic mechanisms in TCMFs. However, there are still shortcomings in semantic modeling, interpretability, and role mapping, and there is an urgent need to combine them with the theoretical knowledge structure of TCM to improve the logical expressiveness and clinical usability of the model.

4.1.3. Knowledge graph

A knowledge graph is a knowledge modeling framework that represents entities and their semantic relationships using a graph structure [164]. It constructs a semantic network of multimodal information in the form of "entity-relationship-entity" triples, effectively integrating heterogeneous data resources such as drugs, targets, pathways, and diseases. KG-MTL combines large-scale biological knowledge graphs with multi-task learning strategies to integrate molecular graphs and biological entity information, thereby enhancing the accuracy of molecular interaction predictions [165]. KGANSynergy constructs knowledge graphs containing relationships such as drug-protein, cell line-protein, and protein-protein, and introduces attention mechanisms to capture complex synergistic interaction pathways [166]. The main advantage of knowledge graphs in modeling molecular compatibility theory in TCM is that they can carry the semantic structure expression of compatibility relationships, which is particularly suitable for constructing the chain logic of active compounds in TCM, functions, and dosages between "monarch, minister, assistant, envoy" and assisting in the discovery of potential synergistic mechanisms through path inference. At the data level, KG can integrate multi-source information such as literature, herbal annotations, and omics experiments to improve the semantic consistency and mechanism interpretability of TCM modeling. However, its construction relies on high-quality, structurally clear data, and the TCM field currently faces challenges such as inconsistent terminology and scarce role annotations, which limit the depth of knowledge representation and model generalization capabilities. Therefore, the promotion of knowledge graphs in TCM theory modeling needs to be combined with the construction of TCM ontologies, natural language information extraction, and multi-perspective knowledge embedding methods to gradually promote the closed-loop integration of "traditional semantics-structural expression-model construction". Therefore, KG shows good potential in constructing the semantic structure of TCMFs and is expected to serve as an important modeling vehicle for integrating knowledge reasoning and mechanism interpretation.

4.2. The analysis of model adaptability

TCM collaborative prediction faces challenges such as data heterogeneity, complex structure, and strong semantic dependence. The application of AI models in this field must simultaneously consider data adaptability, structural expressiveness, and theoretical interpretability. In terms of data representation, traditional models such as SVM and RF require fewer samples and are suitable for preliminary modeling with small samples, but they have limited support for high-dimensional and structural data. DNNs have automatic feature extraction capabilities and are suitable for processing complex data such as molecular structures or pharmacological indicators, but they are sensitive to sample size and distribution. Transformers are suitable for sequence and semantic inputs and have advantages in integrating linguistic features such as pharmacological properties and meridian tropism, but they rely on a large amount of high-quality labeled data. In contrast, GNNs have inherent structural modeling capabilities and are suitable for processing TCM compound-target-disease networks, making them the current mainstream solution. HGNN and KG models perform better in heterogeneous structure representation, but they have a high engineering threshold and rely on complete and standardized knowledge graph support.

In terms of TCMFs structure and TCM semantic modeling, GNN-type models can identify higher-order relationships and role divisions such as "monarch, minister, assistant, envoy" in combinations through graph structures and attention mechanisms. For example, GraphAI for TCM integrates herbal medicine, compounds, targets, and diagnostic information, and embeds semantic nodes such as medicinal properties, medicinal flavors, and meridian tropism, achieving graphical representation of formula structures and quantitative identification of role contributions. It accurately identified highly synergistic drug pairs in the Huashi Baidu formula and validated the mechanisms.

Transformer models excel at extracting semantic features such as "cold-heat" and "ascending-descending" from literature, aiding in the discovery of potential formula relationships and making them more suitable for semantic-structural fusion-assisted modeling. LLMs also demonstrate unique potential in modeling TCM compatibility. For example, models such as OpenTCM [167] and Tianyi [168] can extract cross-context knowledge and perform semantic understanding based on medical literature, classical prescriptions, and texts like *Compendium of Materia Medica*. These capabilities support the construction of modern semantic representations of traditional compatibility relationships.

Furthermore, models like TCMChat [125] and TCM-FTP [169] integrate elements such as medicinal properties, functional roles, dosages, and symptoms. Using attention mechanisms, they generate structured embeddings that enable preliminary modeling of the "monarch, minister, assistant, envoy" framework. In addition, approaches such as OpenTCM and Lingdan [170] combine language models with knowledge graph architectures, enabling multimodal input fusion across text, graphs, and omics data. This integration enhances the model's logical reasoning and interpretability, offering more structured support for the theoretical foundations and clinical value of TCM compatibility.

In terms of model interpretability, the "black box" nature of deep learning models limits the ability to explain the mechanisms behind prediction results. Recent progress indicates that incorporating knowledge graphs [166], attention mechanisms [171], and path visualization tools, such as SHAP [71, 171] and GNNExplainer [172], can effectively enhance model transparency and trustworthiness. GraphAI for TCM uses multi-layer graph structures and attention weight mapping to demonstrate the semantic division of labor and pathway regulation mechanisms of drugs in compound formulas, proving the bridging role of explainable models between theory and mechanisms [173]. Furthermore, explainability should not be limited to feature scoring but should also serve the "holistic" combination logic of TCM. Moving from "who is important" to "how they collaborate" is essential to achieving the deep integration of TCM theory and modern AI models, driving the transformation of models from predictability to understandability.

5. Exploring AI modeling of the "monarch, minister, assistant, envoy" compatibility theory: from empirical stratification to structure-driven

As the core theory of the Chinese medicine combination system, the "monarch, minister, assistant, envoy" compatibility theory is not only a summary of traditional clinical experience, but also reflects the functional division of labor and synergistic relationship in the structure of TCMFs [173]. This theory emphasizes the hierarchical structure of drugs in terms of action intensity, functional objectives, combination synergy, and toxicity and side effect harmonization, which is the "structural logic" unique to Chinese medicine [173]. In recent years, network pharmacology research has gradually restored the functional hierarchy in combinations from three levels [174]: active compounds in TCM, targets, and pathways, as well as the holistic regulation of multiple drugs, multiple targets, and multiple pathways in the human body by TCM. With the increasing involvement of AI in modern TCM research, whether this combination wisdom can be effectively embedded in the model structure not only affects the accuracy of predictions but also determines the model's compatibility and expansion capabilities in TCM theory.

5.1. Expression of "monarch, minister, assistant, envoy" in AI models

Introducing the "monarch, minister, assistant, envoy" structure into AI models is a key step in realizing the integration of Chinese medicine combination modeling and Chinese medicine theory. On the one hand, it is necessary to convert the functional roles of "monarch medicine", "minister medicine", "assistant medicine", and "courier medicine" into recognizable semantic features. on the other hand, the model structure itself must support the participation of these roles in training and prediction, thereby reflecting the hierarchical and synergistic logic of TCMFs.

Existing research has explored the extraction of drug functional roles from databases such as TCMID, ETCM, and HERB, and integrated them with information on drug properties and meridian tropism as node attributes in GNN models [154]. GNNs, especially HGNNs, possess inherent structural modeling capabilities, enabling the representation of compound interactions, functional synergies, or target overlaps between herbs. Through attention mechanisms and edge weight learning, the model can identify the dominant role of the principal drug and the regulatory and guiding functions of the auxiliary drugs, achieving a "structure-function" mapping aligned with TCM principles. Additionally, through natural language processing and other methods, common "monarch, minister, assistant, envoy" structural templates can be extracted from classical formulas, serving as graph structure priors to guide the model in learning traditional combination patterns, thereby enhancing stability in small-sample tasks.

Further research, such as GraphAI for TCM, has extended the "monarch, minister, assistant, envoy" structure modeling to the molecular level. This model constructs a multidimensional knowledge graph (TCM-MKG) of Chinese medicine-compounds-targets-diseases-semantic tags, integrating more than 120 000 compound structures and 2 million target-related data points, and modeling the molecular synergistic strength between drugs through a graph attention mechanism [175]. In the modeling of Huazhi Baidu Tang, astragalus was given the highest attention weight, demonstrating its core status as the "monarch medicine". The model also introduces virtual nodes such as "pharmacological properties" and "meridian tropism" to construct an HGNN structure, thereby strengthening the mapping between TCM semantics and molecular mechanisms. Combined with pathway enrichment analysis, the study revealed that the Huangqi-Reed Root drug pair exerts immune regulatory effects by regulating pathways such as "neuroactive ligand-receptor interactions", forming a "clear hierarchy" in the graph visualization.

5.2. AI modeling of formula changes in classic TCM texts

TCM tailors treatment methods and personalized prescriptions to the specific symptoms of each patient, embodying the important principles of syndrome differentiation and individualized treatment in TCM theory [176-178]. The "monarch, minister, assistant, envoy" structure reflects the functional hierarchy and dosage priorities, embodying the characteristics of personalized treatment in TCM [179–180]. In the theory of TCM compatibility, not only are the types and functions of drugs clearly defined, but the fine adjustment of drug dosage ratios also has an important impact on the efficacy of TCMFs [181]. Even if the types of drugs in a formula remain unchanged, changes in the dosage ratios of the drugs in the formula will result in different pharmacological effects and functions. Take the "*Si Ni Decoction*" and "*Tongmai Si Ni Decoction*" from the *Shang Han Lun* as examples. Both formulas share the same herbal compounds, including *Aconiti Lateralis Radix Praeparata* (Fuzi), *Zingiberis Rhizonma* (Ganjiang), and *Glycyrrhizae Radix Et Rhizoma* (Gancao). However, in the latter formula, the dosage of dried ginger is increased from 1 liang to 3 liang (a TCM unit of weight, approximately 37.3 grams in ancient times), and *Aconiti Lateralis Radix Praeparata* is specifically selected as "large", resulting in a significant adjustment of the dosage ratios [175]. This change not only enhances the "resuscitate yang" (hui yang) effect but also reflects an adaptive shift to address the more severe pathological mechanism of *"yin waxing with yang waning"*. Similarly, in the *"Bai Tong Tang"* and *"Bai Tong Jia Zhu Dan Zhi Tang"* formulas from the *Essential Prescriptions from the Golden Cabinet* (Jin Gui Yao Lue), while retaining the *Aconiti Lateralis Radix Praeparata* and *Zingiberis Rhizonma* combination, the formulas are adjusted to address the complex pathology of "cold-induced diarrhea complicated with dampness" (han li jia shi). Although Li Shizhen did not explicitly specify dosage configurations in the *Compendium of Materia Medica*, he summarized the complementary effects of the Aconite-Ginger combination from a pharmacological perspective through de-

scriptions such as "combined use multiplies efficacy" and "primarily restores yang" emphasizing their synergistic pharmacological basis for "warm yang". These differences indicate that, as a stable drug pair, the dosage ratio and functional weight of *Aconiti Lateralis Radix Praeparata-Zingiberis Rhizonma* are dynamically adjusted based on the pathogenesis, treatment methods, and the physician's therapeutic strategy, reflecting the evolutionary mechanism of the theory of TCMFs.

Now, AI can provide a new perspective for modeling and deducing the patterns of changes in the use of Chinese medicine in terms of the proportions, efficacy, and therapeutic effects of TCMFs. On the one hand, pre-trained language models or LLMs fine-tuned with Chinese medicine knowledge can be used to deeply mine information from classical Chinese medicine literature such as *Treatise on Cold Damage*, *Essential Prescriptions from the Golden Cabinet*, and *Compendium of Materia Medica* [182]. This information includes the relationship between drug combinations, dosage changes, and symptoms [169]. Through these models, it is possible to establish a chain of reasoning between TCMFs, dosage changes, efficacy, and diseases, further explore the potential semantic connections between them, and capture the evolutionary process of combination ideas between different periods and schools of thought. On the other hand, it is also possible to construct a knowledge graph with drugs, formulas, efficacy, causes, and main symptoms as nodes. Then, GNN or hypergraph learning algorithms can be introduced, with dosage ratios as edge weights and functional weights assigned to simulate the impact of dosage variations on efficacy and functionality. Additionally, reinforcement learning offers a new modeling approach [183]. We can view ancient medical practitioners' adjustments to dosage as a strategy optimization process. Different dosages are treated as different actions, the patient's condition as environmental variables, and treatment efficacy and theoretical consistency as reward functions. Through this approach, we can simulate the dynamic optimization process in TCM clinical practice, where treatment plans are continuously adjusted based on efficacy feedback and theoretical consistency, thereby exploring the optimal solutions for dosage, functionality, efficacy, and disease.

6. Discussion

6.1. Data bottleneck

Although the traditional applications and experiential accumulation of TCM provide a foundation for drug combinations, current research on TCM combination synergy still largely relies on experiential inheritance and small-scale experimental validation. In contrast to the high-throughput screening systems used in western medicine for drug combination selection, TCM lacks high-quality, large-scale, and systematic experimental data to support research in this area [80]. Additionally, the quality of multi-omics data in current TCM research is inconsistent, and the comparability of data across different studies is weak [77]. Most existing data come from individualized treatments or case studies in traditional literature, which leads to a lack of standardization and generalizability [184]. This creates a significant challenge in integrating multi-omics data across studies and laboratories. Existing databases, such as TCMSP [94], ETCM, and SymMap [102], although functional to some extent, still face issues such as insufficient data coverage and weak integration. These databases urgently need further improvement. To support the accurate modeling of TCM efficacy compounds synergy prediction, it is necessary to build high-quality databases that cover the chemical composition of TCM, target pathways, pharmacological effects, and omics responses [80]. These databases should not only contain high-quality data but also provide standardized annotations and ensure continuous data updates [182].

6.2. Challenges in the transformation of the connotation of TCM

The compatibility theory of TCM, as a crucial compound of TCM, embodies rich clinical experience and theoretical knowledge. The complexity and systematic nature of these principles largely reflect the essence of TCM. However, transforming these compatibility principles, which are based on traditional experience and intuitive understanding, into modern biological and computational languages presents numerous challenges [185]. Firstly, the complexity of TCM compatibility makes it difficult to precisely describe using the language of modern biology. The mechanisms of action of TCM are not singular or linear but are multi-factorial and multi-dimensional. Modern biology, on the other hand, tends to focus on specific molecular mechanisms, attempting to explain therapeutic effects by studying individual drugs or biological pathways [185]. Secondly, modern biological research generally relies on large-scale experiments and quantitative analysis, often standardizing disease and treatment processes into specific biological markers or therapeutic targets. In contrast, TCM emphasizes individualized treatment based on patient differences, with the efficacy of herbal combinations often varying from person to person [186–187]. This individualized approach conflicts with the standardized, systematic nature of modern biological research [188]. It is precisely this variability that complicates the process of translating TCM compatibility principles into scientific and computational languages. Moreover, the interactions involved in TCM compatibility are often complex and multi-dimensional. This intricate relationship makes it a significant challenge to accurately simulate the effects of TCM combinations using computational languages or algorithmic models [189]. Particularly when considering that the effects of TCM are not due to a single pharmacological effect but rather the combined action of multiple effects [45], it becomes evident that computational languages must possess sufficient flexibility and diversity. However, existing technologies are still struggling to cope with this complexity. Therefore, despite recent advancements in the modernization of TCM compatibility research, how to effectively integrate this traditional knowledge with modern biological and computational languages remains a pressing issue.

6.3. Toward a predictive-validation loop in AI-driven TCM modeling

In the future, it will be necessary to promote the establishment of a closed-loop interdisciplinary mechanism of "prediction-verification-optimization". When constructing models, TCM experts should be deeply involved in the definition of labels and the design of mechanism attribution, clarifying the functions and dosage requirements of "monarch medicine", "minister medicine", "assistant medicine" and "envoy medicine" [190–191]. AI researchers should then design model tasks based on this, such as using multi-layer graph convolutional learning to study role differentiation and identifying action pathways through multi-hop paths, to generate data results with biological mechanism hypotheses. Experimental teams can then conduct *in vitro* or animal model validation based on these results, forming a controlled biological validation feedback loop. Even if current work focuses primarily on computational prediction or pathway analysis, the results should provide mechanism hypotheses that can be translated into experimental designs, as well as clear candidate targets or drug pair structures, to provide concrete support for subsequent validation and model iteration. Research proposals that encompass the entire process of "modeling-mechanism output-experimental validation" should be encouraged. Special funding mechanisms and shared platform construction should be established to support the construction of structural a priori models, multi-center data integration, and experimental feedback iteration. Interdisciplinary integration should not merely be formal

collaboration but should redefine problem-definition methods, logical frameworks, and research objectives, establishing a new paradigm centered on "verifiable predictions" to advance TCM formula theory toward a unified system of structural modeling, mechanism explanation, and clinical practice.

7. Conclusion

Grounded in molecular compatibility theory, this article presents a comprehensive research framework that integrates TCM principles, TCM data, and AI technologies. It systematically delineates key concepts within the molecular compatibility theory, highlighting their relevance to contemporary drug synergy and molecular compatibility modeling. The study further explores the applicability of mainstream AI algorithms in predicting the synergistic effects of active compounds in TCM, while addressing strategies for managing heterogeneous TCM data and improving model interpretability. Emphasis is placed on the potential of AI to simulate the complex structural logic embedded in the "monarch, minister, assistant, envoy" framework. By bridging theoretical foundations with technological tools and data resources, this work seeks to promote the innovative advancement of TCM theory in the era of modern science and technology.

Funding

This work was supported by the National Key Research and Development Program of China (No. 2024YFC3506900), Science and Technology Program of Tianjin (No. 24ZXZSSS00460), Special Project for Technological Innovation in New Productive Forces of Modern Chinese Medicines (No. 24ZXZKSY00010).

Declaration of competing interest

The authors declare that there are no conflicts of interest.

References

1 Dong QF, Bao HL, Wang JG, et al. Liver fibrosis and MAFLD: the exploration of multi-drug combination therapy strategies. *Front Med*. 2023;10:1120621. https://doi.org/10.3389/fmed.2023.1120621.
2 Jia J, Zhu F, Ma XH, et al. Mechanisms of drug combinations: interaction and network perspectives. *Nat Rev Drug Discov*. 2009;8(2):111-128. https://doi.org/10.1038/nrd2683.
3 Hamoud R, Zimmermann S, Reichling J, et al. Synergistic interactions in two-drug and three-drug combinations (thymol, EDTA and vancomycin) against multi drug resistant bacteria including *E. coli*. *Phytomedicine*. 2014;21(4):443-447. https://doi.org/10.1016/j.phymed.2013.10.016.
4 Pool JL. Is it time to move to multidrug combinations? *Am J Hypertens*. 2003;16(S3): 36S-40S. https://doi.org/10.1016/j.amjhyper.2003.07.005.
5 Madani TSA, Soltan GL, Manem VSK, et al. Predictive approaches for drug combination discovery in cancer. *Brief Bioinform*. 2018;19(2):263-276. https://doi.org/10.1093/bib/bbw104.
6 Kong WKX, Midena G, Chen YJ, et al. Systematic review of computational methods for drug combination prediction. *Comput Struct Biotechnol J*. 2022;20:2807-2814. https://doi.org/10.1016/j.csbj.2022.05.055.
7 Tsigelny IF. Artificial intelligence in drug combination therapy. *Brief Bioinform*. 2019;20(4):1434-1448. https://doi.org/10.1093/bib/bby004.
8 Rani P, Dutta K, Kumar V. Artificial intelligence techniques for prediction of drug synergy in malignant diseases: past, present, and future. *Comput Biol Med*. 2022;144:105334. https://doi.org/10.1016/j.compbiomed.2022.105334.
9 Pourmousa M, Jain S, Barnaeva E, et al. AI-driven discovery of synergistic drug combinations against pancreatic cancer. *Nat Commun*. 2025;16(1):4020. https://doi.org/10.1038/s41467-025-56818-6.
10 Peng Z, Ding YL, Zhang PF, et al. Artificial intelligence application for anti-tumor drug synergy prediction. *Curr Med Chem*. 2024;31(40):6572-6585. https://doi.org/10.2174/0109298673290777240301071513.
11 Kumar V, Dogra N. A comprehensive review on deep synergistic drug prediction techniques for cancer. *Arch Comput Methods Eng*. 2022;29(3):1443-1461. https://doi.org/10.1007/s11831-021-09617-3.
12 Wang Y, Wang JJ, Liu Y. Deep learning for predicting synergistic drug combinations: state-of-the-arts and future directions. *Clin Transl Discov*. 2024;4(3):e317. https://doi.org/10.1002/ctd2.317.
13 Pan YC, Ren HT, Lan L, et al. Review of predicting synergistic drug combinations. *Life*. 2023;13(9):1878. https://doi.org/10.3390/life13091878.
14 Abbasi F, Rousu J. New methods for drug synergy prediction: a mini-review. *Curr Opin Struct Biol*. 2024;86:102827. https://doi.org/10.1016/j.sbi.2024.102827.
15 Chen B, Garmire L, Calvisi DF, et al. Harnessing big 'omics' data and AI for drug discovery in hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol*. 2020;17(4):238-251. https://doi.org/10.1038/s41575-020-0288-6.
16 Ali AM, Mohammed MA. A comprehensive review of artificial intelligence approaches in omics data processing: evaluating progress and challenges. *IJMSCS*. 2024;2:114-167. https://doi.org/10.59543/ijmscs.v2i.8703.
17 Ali H. Artificial intelligence in multi-omics data integration: advancing precision medicine, biomarker discovery and genomic-driven disease interventions. *Int J Sci Res Arch*. 2023;8(1):1012-1030. https://doi.org/10.30574/ijsra.2023.8.1.0189.
18 Holzinger A, Haibe-Kains B, Jurisica I. Why imaging data alone is not enough: AI-based integration of imaging, omics, and clinical data. *Eur J Nucl Med Mol Imaging*. 2019;46(13):2722-2730. https://doi.org/10.1007/s00259-019-04382-9.
19 Baciu C, Xu C, Alim M, et al. Artificial intelligence applied to omics data in liver diseases: enhancing clinical predictions. *Front Artif Intell*. 2022;5:1050439. https://doi.org/10.3389/frai.2022.1050439.
20 Wu Y, Xie L. AI-driven multi-omics integration for multi-scale predictive modeling of causal genotype-environment-phenotype relationships. *Comput Struct Biotechnol J*. 2025; 27:265-277. https://doi.org/10.1016/j.csbj.2024.12.030.
21 Tong L, Shi WQ, Isgut M, et al. Integrating multi-omics data with EHR for precision medicine using advanced artificial intelligence. *IEEE Rev Biomed Eng*. 2023;17:80-97. https://doi.org/10.1109/RBME.2023.3324264.
22 Nam Y, Kim J, Jung SH, et al. Harnessing artificial intelligence in multimodal omics data integration: paving the path for the next frontier in precision medicine. *Annu Rev Biomed Data Sci*. 2024;7(1):225-250. https://doi.org/10.1146/annurev-biodatasci-102523-103801.
23 Fawaz A, Ferraresi A, Isidoro C. Systems biology in cancer diagnosis integrating omics technologies and artificial intelligence to support physician decision making. *J Pers Med*. 2023;13(11):1590. https://doi.org/10.3390/jpm13111590.
24 Wang SP, Hu YY, Tan W, et al. Compatibility art of traditional Chinese medicine: from the perspective of herb pairs. *J Ethnopharmacol*. 2012;143(2):412-423. https://doi.org/10.1016/j.jep.2012.07.033.
25 Chen SL, Jiang JG. Application of gene differential expression technology in the mechanism studies of nature product-derived drugs. *Expert Opin Biol Ther*. 2012;12(7):823-839. https://doi.org/10.1517/14712598.2012.683858.
26 Yang Y, Zhang ZQ, Li SP, et al. Synergy effects of herb extracts: pharmacokinetics and pharmacodynamic basis. *Fitoterapia*. 2014;92:133-147. https://doi.org/10.1016/j.fitote.2013.10.010.
27 Yuan H, Ma QQ, Cui HY, et al. How can synergism of traditional medicines benefit from network pharmacology? *Molecules*. 2017;22(7): 1135. https://doi.org/10.3390/molecules22071135.
28 Jia W, Gao WY, Yan YQ, et al. The rediscovery of ancient Chinese herbal formulas. *Phytother Res*. 2004;18(8): 681-686. https://doi.org/10.1002/ptr.1506.
29 Ren X, Yan CX, Zhai RX, et al. Comprehensive survey of target prediction web servers for traditional Chinese medicine. *Heliyon*. 2023;9(8):e19151. https://doi.org/10.1016/j.heliyon.2023.e19151.
30 Liu YS, Li X, Chen C, et al. Exploration of compatibility rules and discovery of active ingredients in TCM formulas by network pharmacology. *Chin Herb Med*. 2024;16(4):572-588. https://doi.org/10.1016/j.chmed.2023.09.008.
31 Luo LF, Zhou JR, Liu XN, et al. Development of modern Chinese medicine guided by molecular compatibility theory. *J Adv Res*. 2024;73:713-728. https://doi.org/10.1016/j.jare.2024.08.005.
32 Zhang GN, Ashby CR, Zhang YK, et al. The reversal of antineoplastic drug resistance in cancer cells by β-elemene. *Chin J Cancer*. 2015;34(11):1-8. https://doi.org/10.1186/s40880-015-0048-0.
33 Zhai BT, Zhang NN, Han XM, et al. Molecular targets of β-elemene, a herbal extract used in traditional Chinese medicine, and its potential role in cancer therapy: a review. *Biomed Pharmacother*. 2019;114:108812. https://doi.org/10.1016/j.biopha.2019.108812.
34 Jiang XY, Shi LP, Zhu JL, et al. Elemene antitumor drugs development based on "molecular compatibility theory" and clinical application: a retrospective and prospective outlook. *Chin J Integr Med*. 2024;30(1):62-74. https://doi.org/10.1007/s11655-023-3714-0.
35 Xie T, Wang SL, Zeng ZW, et al. Molecular Compatibility Theory of Modern Chinese Medicine and Systems Biology. Proceedings of the 2014 National Academic Symposium on Traditional Chinese Medicine.2014;61-64.
36 Wang SL, Xie T, Sun M, et al. Theory and practice of molecule compatibility. *Chin J Exp Tradit Med Formulae*. 2010;16(15):222-224. https://doi.org/10.13422/j.cnki.syfjx.2010.15.014.
37 Dong QQ, Wang SW. Molecular Chinese materia medica and molecular compatibility theory. *Electron J Clin Med Lit*. 2014;1(2): 124-127. https://doi.org/10.16281/j.cnki.jocml.2014.02.021.
38 Xie T. Elemene antitumor drugs: molecular compatibility theory and its applications in new drug development and clinical practice. 1st ed. Amsterdam: Elsevier; 2022.
39 Lu LK, Lu TS, Tian CY, et al. AI: bridging ancient wisdom and modern innovation in traditional Chinese medicine. *JMIR Med Inform*. 2024;12(1):e58491. https://doi.org/10.2196/58491.
40 Li Y, Liu XJ, Zhou JW, et al. Artificial intelligence in traditional Chinese medicine: advances in multi-metabolite multi-target interaction modeling. *Front Pharmacol*. 2025;16:1541509. https://doi.org/10.3389/fphar.2025.1541509.
41 Yao Y, Zhang XD, Wang ZZ, et al. Deciphering the combination principles of traditional Chinese medicine from a systems pharmacology perspective based on Ma-huang Decoction. *J Ethnopharmacol*. 2013;150(2):619-638.

https://doi.org/10.1016/j.jep.2013.09.018.
42 Zhou X, Seto SW, Chang D, et al. Synergistic effects of Chinese herbal medicine: a comprehensive review of methodology and current research. *Front Pharmacol*. 2016;7:201. https://doi.org/10.3389/fphar.2016.00201.
43 Luan X, Zhang LJ, Li XQ, et al. Compound-based Chinese medicine formula: from discovery to compatibility mechanism. *J Ethnopharmacol*. 2020;254: 112687. https://doi.org/10.1016/j.jep.2020.112687.
44 Ma SS, Jia CH, Guo J. The metaphor analysis of "monarch, minister, assistant and envoy" in TCM formula based on the concept of "a prescription being astate". *J Beijing Univ Tradit Chin Med*. 2019;42(02):93-98. https://doi.org/10.3969/j.issn.1006-2157.2019.02.001.
45 Che CT, Wang ZJ, Chow MSS, et al. Herb-herb combination for therapeutic enhancement and advancement: theory, practice and future perspectives. *Molecules*. 2013;18(5):5125-5141. https://doi.org/10.3390/molecules18055125.
46 Ma L, Zhang X, Xu XM, et al. Compatibility principle in the Tanyu Tongzhi Formula revealed by a cell-based analysis. *J Ethnopharmacol*. 2019;231:507-515. https://doi.org/10.1016/j.jep.2018.11.043.
47 Li S, Zhang B. Traditional Chinese medicine network pharmacology: theory, methodology and application. *Chin J Nat Med*. 2013;11(2):110-120. https://doi.org/10.1016/S1875-5364(13)60037-0.
48 Wang AY, Peng HY, Wang YD, et al. NP-TCMtarget: a network pharmacology platform for exploring mechanisms of action of traditional Chinese medicine. *Brief Bioinform*. 2025;26(1):bbaf078. https://doi.org/10.1093/bib/bbaf078.
49 Zhang WZ, Mao XQ, Sun XN, et al. Research progress on molecular compatibility in integrated Chinese and Western medicine to foster synergy and reverse resistance of cancer cells to anticancer drugs. *Chin J Clin Oncol*. 2021;48(11):566-570. https://doi.org/10.3969/j.issn.1000-8179.2021.11.371.
50 Liu J, Hu XJ, Jin B, et al. β-Elemene induces apoptosis as well as protective autophagy in human non-small-cell lung cancer A549 cells. *J Pharm Pharmacol*. 2012;64(1):146-153. https://doi.org/10.1111/j.2042-7158.2011.01371.x.
51 Cai SZ, Xiong QW, Zhao LN, et al. *β*-Elemene triggers ROS-dependent apoptosis in glioblastoma cells through suppressing STAT3 signaling pathway. *Pathol Oncol Res*. 2021;27:594299. https://doi.org/10.3389/pore.2021.594299.
52 Qureshi MZ, Attar R, Romero MA, et al. Regulation of signaling pathways by *β*-elemene in cancer progression and metastasis. *J Cell Biochem*. 2019;120(8):12091-12100. https://doi.org/10.1002/jcb.28624.
53 Jiang ZY, Jacob JA, Loganathachetti DS, et al. β-Elemene: mechanistic studies on cancer cell interaction and its chemosensitization effect. Front. *Pharmacol*. 2017;8:105. https://doi.org/10.3389/fphar.2017.00105.
54 Fang YY Kang YH, Zou H, et al. *β*-elemene attenuates macrophage activation and proinflammatory factor production *via* crosstalk with Wnt/*β*-catenin signaling pathway. *Fitoterapia*. 2018;124:92-102. https://doi.org/10.1016/j.fitote.2017.10.015.
55 Liu XL, Liu HM. Clinical efficacy and safety of Elemene Injection combined with chemotherapy in the treatment of advanced lung cancer. *J Aerosp Med*. 2017;28(6):720-722. https://doi.org/10.3969/j.issn.2095-1434.2017.06.039.
56 Zhai BT, Zeng YY, Zeng ZW, et al. Drug delivery systems for elemene, its main active ingredient *β*-elemene, and its derivatives in cancer therapy. *Int J Nanomed*. 2018;13:6279-6296. https://doi.org/10.2147/IJN.S174527.
57 Yan MM, Jiao LM, Feng WZ. Observation on the application effects of *β*-elemene adjuvant chemotherapy after lung adenocarcinoma surgery. *Pract J Clin Integr Tradit Chin West Med*. 2025;25(12):11-14. https://doi.org/10.13638/j.issn.1671-4040.2025.12.004.
58 Wang HY, Ma YY. *β*-Elemene alleviates cisplatin resistance in oral squamous cell carcinoma cell *via* inhibiting JAK2/STAT3 pathway *in vitro* and *in vivo*. *Cancer Cell Int*. 2022;22(1):244. https://doi.org/10.1186/s12935-022-02650-7.
59 Wang L, Zhou GB, Liu P, et al. Dissection of mechanisms of Chinese medicinal formula Realgar-Indigo naturalis as an effective treatment for promyelocytic leukemia. *PNAS*. 2008;105(12):4826-4831. https://doi.org/10.1073/pnas.0712365105.
60 Huang DP, Yang LC, Chen YQ, et al. Long-term outcome of children with acute promyelocytic leukemia: a randomized study of oral versus intravenous arsenic by SCCLG-APL group. *Blood Cancer J*. 2023;13(1):178. https://doi.org/10.1038/s41408-023-00949-w.
61 Chen SY, Chen ZJ Wang Y, et al. Targeted delivery of Chinese herb pair-based berberine/tannin acid self-assemblies for the treatment of ulcerative colitis. *J Adv Res*. 2022;40:263-276. https://doi.org/10.1016/j.jare.2021.11.017.
62 Ji S, He DD, Su ZY, et al. P450 enzymes-based metabolic interactions between monarch drugs and the other constituent herbs: a strategy to explore compatibility mechanism of Sangju-Yin. *Phytomedicine*. 2019;58: 152866. https://doi.org/10.1016/j.phymed.2019.152866.
63 Li X, Liu ZQ, Liao J, et al. Network pharmacology approaches for research of traditional Chinese medicines. *Chin J Nat Med*. 2023;21(5):323-332. https://doi.org/10.1016/S1875-5364(23)60429-7.
64 Tang YP, Xu DQ, Yue SJ, et al. Modern research thoughts and methods on bio-active components of TCM formulae. *Chin J Nat Med*. 2022;20(7):481-493. https://doi.org/10.1016/S1875-5364(22)60206-1.
65 Wu T, Lin RM, Cui PD, et al. Deep learning-based drug screening for the discovery of potential therapeutic agents for Alzheimer's disease. *J Pharm Anal*. 2024;14(10):101022. https://doi.org/10.1016/j.jpha.2024.101022.
66 Bliss CI. The toxicity of poisons applied jointly 1. *Ann Appl Biol*. 1939;26(3): 585-615. https://doi.org/10.1111/j.1744-7348.1939.tb06990.x.
67 Loewe S. The problem of synergism and antagonism of combined drugs. *Arzneim-Forsch*. 1953;3(6):285-290.
68 Wu LL, Wen YQ, Leng DJ, et al. Machine learning methods, databases and tools for drug combination prediction. *Brief Bioinform*. 2022;23(1):bbab355. https://doi.org/10.1093/bib/bbab355.
69 Preuer K, Lewis RP, Hochreiter S, et al. DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics*. 2018;34(9):1538-1546. https://doi.org/10.1093/bioinformatics/btx806.
70 Kuru HI, Tastan O, Cicek AE. MatchMaker: a deep learning framework for drug synergy prediction. *IEEE/ACM Trans Comput Biol Bioinf*. 2021;19(4): 2334-2344. https://doi.org/10.1109/TCBB.2021.3086702.
71 Liu Q, Xie L. TranSynergy: mechanism-driven interpretable deep neural network for the synergistic prediction and pathway deconvolution of drug combinations. *PLoS Comput Biol*. 2021;17(2):e1008653. https://doi.org/10.1371/journal.pcbi.1008653.
72 Cokol-Cakmak M, Cetiner S, Erdem N, et al. Guided screen for synergistic three-drug combinations. *PLoS One*. 2020;15(7):e0235929. https://doi.org/10.1371/journal.pone.0235929.
73 She SN, Chen HW, Ji W, et al. Deep learning-based multi-drug synergy prediction model for individually tailored anti-cancer therapies. *Front Pharmacol*. 2022;13:1032875. https://doi.org/10.3389/fphar.2022.1032875.
74 Liu JL, Liu JJ, Shen FX, et al. Systems pharmacology analysis of synergy of TCM: an example using saffron formula. *Sci Rep*. 2018;8(1):380. https://doi.org/10.1038/s41598-017-18764-2.
75 Zhou E, Shen Q, Hou Y. Integrating artificial intelligence into the modernization of traditional Chinese medicine industry: a review. *Front Pharmacol*. 2024;15:1181183. https://doi.org/10.3389/fphar.2024.1181183.
76 Li MY, Zhang J. A focus on harnessing big data and artificial intelligence: revolutionizing drug discovery from traditional Chinese medicine sources. *Chem Sci*. 2023;14(39):10628-10630. https://doi.org/10.1039/D3SC90185H.
77 Wang YY, Liu MX, Jafari M, et al. A critical assessment of traditional Chinese medicine databases as a source for drug discovery. *Front Pharmacol*. 2024;15:1303693. https://doi.org/10.3389/fphar.2024.1303693.
78 Fan MY, Jin C, Li DP, et al. Multi-level advances in databases related to systems pharmacology in traditional Chinese medicine: a 60-year review. *Front Pharmacol*. 2023;14:1289901. https://doi.org/10.3389/fphar.2023.1289901.
79 Guo R, Luo XL, Liu JJ, et al. Omics strategies decipher therapeutic discoveries of traditional Chinese medicine against different diseases at multiple layers molecular-level. *Pharmacol Res*. 2020;152:104627. https://doi.org/10.1016/j.phrs.2020.104627.
80 Wen XH, Wang YR, Su C, et al. Integrating multi-omics technologies with traditional Chinese medicine to enhance cancer research and treatment. *QJM*. 2025;hcaf103. https://doi.org/10.1093/qjmed/hcaf103.
81 Shi YL, Liu J, Guan S, et al. Syn-COM: a multi-level predictive synergy framework for innovative drug combinations. *Pharmaceuticals*. 2024;17(9): 1230. https://doi.org/10.3390/ph17091230.
82 Bhat GR, Sethi I, Rah B, et al. Innovative in silico approaches for characterization of genes and proteins. *Front Genet*. 2022;13:865182. https://doi.org/10.3389/fgene.2022.865182.
83 Feehley T, O'Donnell CW, Mendlein J, et al. Drugging the epigenome in the age of precision medicine. *Clin Epigenetics*. 2023;15(1):6. https://doi.org/10.1186/13148-022-01419-z.
84 Cui K, Wu WW, Diao QY. Application and research progress on transcriptomics. *Biotechnol Bull*. 2019;35(7):1-9. https://doi.org/10.13560/j.cnki.biotech.bull.1985.2019-0374.
85 Zhang QP, Huang QJ, Cheng ZP, et al. Exploring the mechanism of Xiaoaiping Injection inhibiting autophagy in prostate cancer based on proteomics. *Chin J Nat Med*. 2025;23(1):64-76. https://doi.org/10.1016/S1875-5364(25)60804-1.
86 Xu MD, Zhao XW, Wang JY, et al. DFFNDDS: prediction of synergistic drug combinations with dual feature fusion networks. *J Cheminf*. 2023;15(1):33. https://doi.org/10.1186/s13321-023-00690-3.
87 Wang JC, Tian CY, Zhang F, et al. Application of metabolomics in the study of active components and mechanism of action of traditional Chinese medicine. *Chin J Ethnomed Ethnopharm*. 2024;33(14):76-80. https://doi.org/10.3969/j.issn.1007-8517.2024.14.rqrmjnyz202414014.
88 Zhang LQ, Qin YF, Chen M. SMILESynergy: anticancer drug synergy prediction based on transformer pre-trained model. *J Biomed Eng*. 2023;40(3):544-551. https://doi.org/10.7507/1001-5515.202209043.
89 Eckhart L, Lenhof K, Herrmann L, et al. How to predict effective drug combinations–moving beyond synergy scores. *iScience*. 2025;28(6):112622. https://doi.org/10.1016/j.isci.2025.112622.
90 Bao F, Li L, Hammerlindl H, et al. Transitive prediction of small-molecule function through alignment of high-content screening resources. *Nat Biotechnol*. 2025;1-11. https://doi.org/10.1038/s41587-025-02729-2.
91 Li ZG, Wu Q, Xing YR. Key concepts in traditional Chinese medicine II. Singapore: Palgrave Macmillan; 2021.
92 Liu BY, Zhou XZ, Wang YH, et al. Data processing and analysis in real-world traditional Chinese medicine clinical data: challenges and approaches. *Stat Med*. 2012;31(7):653-660. https://doi.org/10.1002/sim.4417.
93 Long H, Zhu Y, Jia LR, et al. An ontological framework for the formalization, organization and usage of TCM-Knowledge. *BMC Med Inf Decis Making*. 2019;19(Suppl 2):53. https://doi.org/10.1186/s12911-019-0760-9.
94 Ru JL, Li P, Wang JN, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminf*. 2014;6:1-6. https://doi.org/10.1186/1758-2946-6-13.
95 KongR X, Liu C, Zhang ZZ, et al. BATMAN-TCM 2.0: an enhanced integrative database for known and predicted interactions between traditional Chinese medicine ingredients and target proteins. *Nucleic Acids Res*. 2024;52(D1): D1110-D1120. https://doi.org/10.1093/nar/gkad926.

96 Zhang YQ, Li X, Shi YL, et al. ETCM v2.0: an update with comprehensive resource and rich annotations for traditional Chinese medicine. *Acta Pharm Sin B*. 2023;13(6):2559-2571. https://doi.org/10.1016/j.apsb.2023.03.012.

97 Fang SS, Dong L, Liu L, et al. HERB: a high-throughput experiment-and reference-guided database of traditional Chinese medicine. *Nucleic Acids Res*. 2021;49(D1):D1197-D1206. https://doi.org/10.1093/nar/gkaa1063.

98 Zhang RZ, Yu SJ, Bai H, et al. TCM-Mesh: the database and analytical system for network pharmacology analysis for TCM preparations. *Sci Rep*. 2017; 7(1):2821. https://doi.org/10.1038/s41598-017-03039-7.

99 Zhu X, Yao Q, Yang PS, et al. Multi-omics approaches for in-depth understanding of therapeutic mechanism for traditional Chinese medicine. *Front Pharmacol*. 2022;13:1031051. https://doi.org/10.3389/fphar.2022.1031051.

100 Huang L, Xie DL, Yu YR, et al. TCMID 2.0: a comprehensive resource for TCM. *NAR*. 2018;46(D1):D1117-D1120. https://doi.org/10.1093/nar/gkx1028.

101 Lv QJ, Chen GX, He HH, et al. TCMBank: bridges between the largest herbal medicines, chemical ingredients, target proteins, and associated diseases with intelligence text mining. *Chem sci*. 2023;14(39):10684-10701. https://doi.org/10.1039/D3SC02139D.

102 Wu Y, Zhang FL, Yang K, et al. SymMap: an integrative database of traditional Chinese medicine enhanced by symptom mapping. *Nucleic Acids Res*. 2019;47(D1):D1110-D1117. https://doi.org/10.1093/nar/gky1021.

103 Yan DY, Zheng GH, Wang CC, et al. HIT 2.0: an enhanced platform for herbal ingredients' targets. *Nucleic Acids Res*. 2022;50(D1):D1238-D1243. https://doi.org/10.1093/nar/gkab1011.

104 Knox C, Wilson M, Klinger CM, et al. DrugBank 6.0: the DrugBank knowledgebase for 2024. *Nucleic Acids Res*. 2024;52(D1):D1265-D1275. https://doi.org/10.1093/nar/gkad976.

105 Irwin JJ, Shoichet BK. ZINC–a free database of commercially available compounds for virtual screening. *J Chem Inf Model*. 2005;45(1):177-182. https://doi.org/10.1021/ci049714 +.

106 Wang YL, Xiao JW, Suzek TO, et al. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res*. 2009;37 (suppl_2):W623-W633. https://doi.org/10.1093/nar/gkp456.

107 Szklarczyk D, Santos A, Von Mering C, et al. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res*. 2016;44(D1):D380-D384. https://doi.org/10.1093/nar/gkv1277.

108 Liu TQ, Lin Y, Wen X, et al. BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res*. 2007;35(suppl_1):D198-D201. https://doi.org/10.1093/nar/gkl999.

109 Consortium U. UniProt: a hub for protein information. *Nucleic Acids Res*. 2014;43(Database issue):D204-D212. https://doi.org/10.1093/nar/gku989.

110 Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43(D1):D447-D452. https://doi.org/10.1093/nar/gku1003.

111 Zanzoni A, Montecchi-Palazzi L, Quondam M, et al. MINT: a molecular INTeraction database. *FEBS lett*. 2002;513(1):135-140. https://doi.org/10.1016/s0014-5793(01)03293-8.

112 Hermjakob H, Montecchi-Palazzi L, Lewington C, et al. IntAct: an open source molecular interaction database. *Nucleic Acids Res*. 2004;32(suppl_1): D452-D455. https://doi.org/10.1093/nar/gkh052.

113 Oughtred R, Rust J, Chang C, et al. The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci*. 2021;30(1):187-200. https://doi.org/10.1002/pro.3978.

114 Piñero J, Bravo À, Queralt-Rosinach N, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res*. 2016;45(D1):D833-D839. https://doi.org/10.1093/nar/gkw943.

115 Davis AP, Wiegers TC, Johnson RJ, et al. Comparative toxicogenomics database (CTD): update 2023. *Nucleic Acids Res*. 2023;51(D1):D1257-D1262. https://doi.org/10.1093/nar/gkac833.

116 Yang W, Soares J, Greninger P, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2012;41(D1):D955-D961. https://doi.org/10.1093/nar/gks1111.

117 Ghandi M, Huang FW, Jané-Valbuena J, et al. Next-generation characterization of the cancer cell line encyclopedia. *Nature*. 2019;569 (7757):503-508. https://doi.org/10.1038/s41586-019-1186-3.

118 Forbes S, Clements J, Dawson E, et al. COSMIC 2005. *Br J cancer*. 2006;94(2): 318-322. https://doi.org/10.1038/sj.bjc.6602928.

119 Zhang JJ, Bajari R, Andric D, et at. The international cancer genome consortium data portal. *Nat Biotechnol*. 2019;37(4):367-369. https://doi.org/10.1038/s41587-019-0055-9.

120 Tomczak K, Czerwińska P, Wiznerowicz M. Review the cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. 2015;19(1A):A68-A77. https://doi.org/10.5114/wo.2014.47136.

121 Diamant I, Clarke DJ, Evangelista JE, et al. Harmonizome 3.0: integrated knowledge about genes and proteins from diverse multi-omics resources. *Nucleic Acids Res*. 2025;53(D1):D1016-D1028. https://doi.org/10.1093/nar/gkae1080.

122 Yang ZW, Kotoge R, Piao XH, et al. MLomics: cancer multi-omics database for machine learning. *Sci Data*. 2025;12(1):1-9. https://doi.org/10.1038/s41597-025-05235-x.

123 Li YY, Zhou LW, Qian FC, et al. scImmOmics: a manually curated resource of single-cell multi-omics immune data. *Nucleic Acids Res*. 2025;53(D1):D1162-D1172. https://doi.org/10.1093/nar/gkae985.

124 Yang GX, Liu XH, Shi JY, et al. TCM-GPT: efficient pre-training of large language models for domain adaptation in traditional Chinese medicine. *Comput Methods Programs Biomed Update*. 2024;6:100158. https://doi.org/10.1016/j.cmpbup.2024.100158.

125 Dai YZ, Shao X, Zhang JL, et al. TCMChat: a generative large language model for traditional Chinese medicine. *Pharmacol Res*. 2024;210:107530. https://doi.org/10.1016/j.phrs.2024.107530.

126 Yang S, Zhao H, Zhu S, et al. Zhongjing: enhancing the Chinese medical capabilities of large language model through expert feedback and real-world multi-turn dialogue. AAAI'24/IAAI'24/EAAI'24: Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence. 2024;19368-19376. https://doi.org/10.1609/aaai.v38i17.29907.

127 Chen Z, Wang H, Li CX, et al. Large language models in traditional Chinese medicine: a systematic review. *Acupunct Herb Med*. 2025;5(1):57-67. https://doi.org/10.1097/HM9.0000000000000143.

128 Zhang L, Dong X, Jia S, et al. Node2Vec-DGI-EL: a hierarchical graph representation learning model for ingredient-disease association prediction. *arXiv preprint. 250500236* https://doi.org/10.48550/arXiv.2505.00236.

129 Tanvir F, Saifuddin KM, Hossain T, et al. HeTriNet: Heterogeneous graph triplet attention network for drug-target-disease interaction. *ICDEW: Proceedings of the 2023 IEEE 39th International Conference on Data Engineering Workshops*. 2023;1503-1516. https://doi.org/10.1109/ICDE55515.2023.00119.

130 Li WJ, Ma WJ, Yang MY, et al. Drug repurposing based on the DTD-GNN graph neural network: revealing the relationships among drugs, targets and diseases. *BMC Genomics*. 2024;25(1):584. https://doi.org/10.1186/s12864-024-10499-5.

131 Zheng X, Wu H, Jin H, et al. FMCHS: advancing traditional Chinese medicine herb recommendation with fusion of multiscale correlations of herbs and symptoms. *arXiv preprint. 250305167*. https://doi.org/10.48550/arXiv.2503.

132 Cui YD, Gao B, Liu LH, et al. AMFormulaS: an intelligent retrieval system for traditional Chinese medicine formulas. *BMC Med Inform Decis Mak*. 2021;21: 56. https://doi.org/10.1186/s12911-021-01419-8.

133 Foucquier J, Guedj M. Analysis of drug combinations: current methodological landscape. *Pharmacol Res Perspect*. 2015;3(3):e00149. https://doi.org/10.1002/prp2.149.

134 Yadav B, Wennerberg K, Aittokallio T, et al. Searching for drug synergy in complex dose–response landscapes using an interaction potency model. *Comp Struct Biotechnol J*. 2015;13:504-513. https://doi.org/10.1016/j.csbj.2015.09.001.

135 Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006;24(12): 1565-1567. https://doi.org/10.1038/nbt1206-1565.

136 Wang TY, Chen L, Zhao X. Prediction of drug combinations with a network embedding method. *Comb Chem High Throughput Screen*. 2018;21(10):789-797. https://doi.org/10.2174/1386207322666181226170140.

137 Xu Q, Xiong Y, Dai H, et al. PDC-SGB: prediction of effective drug combinations using a stochastic gradient boosting algorithm. *J Theor Biol*. 2017;417:1-7. https://doi.org/10.1016/j.jtbi.2017.01.019.

138 Sun JJ, Ni QH, Jiang FY, et al. Discovery and validation of traditional Chinese and western medicine combination antirheumatoid arthritis drugs based on machine learning (random forest model). *Biomed Res Int*. 2023;2023(1): 6086388. https://doi.org/10.1155/2023/6086388.

139 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. https://doi.org/10.1038/nature14539.

140 Gawehn E, Hiss JA, Schneider G. Deep learning in drug discovery. *Mol Inform*. 2016;35(1):3-14. https://doi.org/10.1002/minf.201501008.

141 Zhang T, Zhang L, Payne PR, et al. Synergistic drug combination prediction by integrating multiomics data in deep learning models. *Methods Mol Biol*. 2021;2194:223-238. https://doi.org/10.1007/978-1-0716-0849-4_12.

142 Movahedi F, Coyle JL, Sejdić E. Deep belief networks for electroencephalography: a review of recent contributions and future outlooks. *IEEE J Biomed Health Inform*. 2017;22(3):642-652. https://doi.org/10.1109/JBHI.2017.2727218.

143 Chen GC, Tsoi A, Xu H, et al. Predict effective drug combination by deep belief network and ontology fingerprints. *J Biomed Inform*. 2018;85:149-154. https://doi.org/10.1016/j.jbi.2018.07.024.

144 Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw*. 2015;61:85-117. https://doi.org/10.1016/j.neunet.2014.09.003.

145 Jain S, Srivastava R. Multi-modality NDE fusion using encoder–decoder networks for identify multiple neurological disorders from EEG signals. *Technol Health Care*. 2025; 33(5): 2431-2451.https://doi.org/10.1177/09287329241291334.

146 Zhang P, Tu SK, Zhang W, et al. Predicting cell line-specific synergistic drug combinations through a relational graph convolutional network with attention mechanism. *Brief Bioinform*. 2022;23(6):bbac403. https://doi.org/10.1093/bib/bbac403.

147 Kumar N, Srivastava R. Deep learning in structural bioinformatics: current applications and future perspectives. *Brief Bioinform*. 2024;25(3): bbae042. https://doi.org/10.1093/bib/bbae042.

148 Dwivedi VP, Schlegel V, Liu AT, et al. Representation learning of structured data for medical foundation models. *arXiv preprint. 241013351*. https://doi.org/10.48550/arXiv.2410.13351.

149 Parmar N, Vaswani A, Uszkoreit J, et al. Image transformer. *ICML'18: Proceedings of the 35th International Conference on Machine Learning*. 2018; 4055-4064.

150 Jiang J, Chen L, Ke L, et al. A review of transformers in drug discovery and beyond. *J Pharm Anal*. 2024;15(6): 101081. https://doi.org/10.1016/j.jpha.2024.101081.

151 Edwards C, Naik A, Khot T, et al. SynerGPT: in-context learning for personalized drug synergy prediction and drug design. COLM'24: Proceedings of the First Conference on Language Modeling. 2024. https://openreview.net/forum?id=Aaz6R4Tlwv.

152 Liu TY, Chu TY, Luo X, et al. Building a unified model for drug synergy analysis powered by large language models. *Nat Commun*. 2025;16: 4537. https://doi.org/10.1038/s41467-025-59822-y.

153 Corso G, Stark H, Jegelka S, et al. Graph neural networks. *Nat Rev Method Prim*. 2024;4(1): 17. https://doi.org/10.1038/s43586-024-00294-7.

154 Gu LM, Ma YN, Liu SJ, et al. Prediction of herbal compatibility for colorectal adenoma treatment based on graph neural networks. *Chin Med*. 2025;20(1): 1-12. https://doi.org/10.1186/s13020-025-01082-5.

155 Samek W, Montavon G, Lapuschkin S, et al. Explaining deep neural networks and beyond: a review of methods and applications. *Proc IEEE* 2021;109(3): 247-278. https://doi.org/10.1109/JPROC.2021.3060483

156 Zhang ZH, Chen LF, Zhong FS, et al. Graph neural network approaches for drug-target interactions. *Curr Opin Struct Biol*. 2022;73: 102327. https://doi.org/10.1016/j.sbi.2021.102327.

157 Veličković P, Cucurull G, Casanova A, et al. Graph attention networks. ICLR'18: Proceedings of the International Conference on Learning Representations. 2018. https://openreview.net/forum?id=rJXMpikCZ.

158 Wang J, Liu X, Shen S, et al. DeepDDS: deep graph neural network with attention mechanism to predict synergistic drug combinations. *Brief Bioinform*. 2022;23(1): bbab390. https://doi.org/10.1093/bib/bbab390.

159 Li XL, Shen BH, Feng F, et al. Dual-view jointly learning improves personalized drug synergy prediction. *Bioinformatics*. 2024;40(10): btae604. https://doi.org/10.1093/bioinformatics/btae604.

160 Li HJ, Zou L, Kowah JA, et al. Predicting drug synergy and discovering new drug combinations based on a graph autoencoder and convolutional neural network. *Interdiscip Sci*. 2023;15(2): 316-330. https://doi.org/10.1007/s12539-023-00558-y.

161 Hamilton W, Ying Z, Leskovec J. Inductive representation learning on large graphs. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017; 1025-1035. https://doi.org/10.5555/3294771.3294869.

162 Liu X, Song CZ, Liu SC, et al. Multi-way relation-enhanced hypergraph representation learning for anti-cancer drug synergy prediction. *Bioinformatics*. 2022;38(20): 4782-4789. https://doi.org/10.1093/bioinformatics/btac579.

163 Li L, Lü GD, Zheng CH, et al. MHCLSyn: multi-view hypergraph contrastive learning for synergistic drug combination prediction. *Big Data Min Anal*. 2024;7(4): 1273-1286. https://doi.org/10.26599/BDMA.2024.9020054.

164 Hogan A, Blomqvist E, Cochez M, et al. Knowledge graphs. *ACM Comput Surv*. 2021;54(4): 1-37. https://doi.org/10.1145/3447772.

165 Ma T, Lin X, Song B, et al. Kg-mtl: knowledge graph enhanced multi-task learning for molecular interaction. *IEEE Trans Knowl Data Eng*. 2022;35(7): 7068-7081. https://doi.org/10.1109/TKDE.2022.3188154.

166 Zhang G, Gao ZJ, Yan CK, et al. KGANSynergy: knowledge graph attention network for drug synergy prediction. *Brief Bioinform*. 2023;24(3): bbad167. https://doi.org/10.1093/bib/bbad167.

167 He J, Guo Y, Lam L K, et al. Opentcm: a graphrag-empowered llm-based system for traditional Chinese medicine knowledge retrieval and diagnosis. *arXiv preprint. 250420118*. https://doi.org/10.48550/arXiv.2504.20118.

168 Liu Z, Yang T, Wang J, et al. Tianyi: A traditional Chinese medicine all-rounder language model and its real-world clinical practice. *Inf Fusion*. 2026;126:103663. https://doi.org/10.1016/j.inffus.2025.103663.

169 Zhou XZ, Dong X, Li CH, et al. TCM-FTP: fine-tuning large language models for herbal prescription prediction. *2024 IEEE International Conference on Bioinformatics and Biomedicine*. 2024; 4092-4097. https://doi.org/10.1109/BIBM62325.2024.10822451.

170 Hua R, Dong X, Wei Y, et al. Lingdan: enhancing encoding of traditional Chinese medicine knowledge for clinical reasoning tasks with large language models. *J Am Med Inform Assoc*. 2024;31(9): 2019-2029. https://doi.org/10.1093/jamia/ocae087.

171 Tang YC, Gottlieb A. SynPathy: predicting drug synergy through drug-associated pathways using deep learning. *Mol Cancer Res*. 2022;20(5): 762-769. https://doi.org/10.1158/1541-7786.MCR-21-0735.

172 Ying Z, Bourgeois D, You J, et al. Gnnexplainer: generating explanations for graph neural networks. *Adv Neural Inf Process Syst*.2019;32:9240-9251.

173 Li Y, Wang JH, Lin F, et al. A methodology for cancer therapeutics by systems pharmacology-based analysis: a case study on breast cancer-related traditional Chinese medicines. *PLoS One*. 2017;12(1): e0169363. https://doi.org/10.1371/journal.pone.0169363

174 Zhai Y, Liu L, Zhang F, et al. Network pharmacology: a crucial approach in traditional Chinese medicine research. *Chin Med*. 2025;20(1): 8. https://doi.org/10.1186/s13020-024-01056-z.

175 Zeng JQ, Jia XB. Quantifying compatibility mechanisms in traditional Chinese medicine with interpretable graph neural networks. *J Pharm Anal*. 2025;101342. https://doi.org/10.1016/j.jpha.2025.101342.

176 Xu H. Building a new system of TCM syndrome differentiation and treatment of "Formula-Disease-Syndrome". *Chin J Integr Tradit West Med*. 2024;44(12): 1503-1506. https://doi.org/10.13193/j.issn.1673-7717.2019.05.017

177 Liu Q, Liu L, Fan Y. Strategies and approaches on exploring compatibilities of Chinese herbs and prescriptions. *Chin Arch Tradit Chin Med*. 2019;37(5): 1092-1095. https://doi.org/10.13193/j.issn.1673-7717.2019.05.017.

178 Zheng YX, Wang YY, Wang L, et al. Modern research and thinking on compatibility mechanism of reducing toxicity of traditional Chinese medicine. *Chin Tradit Herb Drugs*. 2023;54(2): 386-395. https://doi.org/10.7501/j.issn.2357-2020.0005

179 Steven LCT, Yi GX. Discussion on relevance and studies of prescription compatibility in Chinese medicine. *Chin J Integr Med*. 2021;27(10):788-793. https://doi.org/10.1007/s11655-020-3217-1.

180 Zhou MM, Hong YL, Lin X, et al. Recent pharmaceutical evidence on the compatibility rationality of traditional Chinese medicine. *J Ethnopharmacol*. 2017;206: 363-375. https://doi.org/10.1016/j.jep.2017.06.007.

181 Wu WH, Sun HL. On the variation law of the proportion of dosages in prescriptions of treatise on febrile and miscellaneous diseases. *Acta Chin Med Pharmacol*. 1994;(1): 12-14. https://doi.org/10.19664/j.cnki.1002-2392.1994.01.005.

182 Song Z, Chen G, Chen CY. AI empowering traditional Chinese medicine? *Chem Sci*. 2024;15(41): 16844-16886. https://doi.org/10.1039/D4SC04107K.

183 Yang K, Yu ZC, Su X, et al. PrescDRL: deep reinforcement learning for herbal prescription planning in treatment of chronic diseases. *Chin Med*. 2024;19(1): 144. https://doi.org/10.1186/s13020-024-01005-w.

184 Yang Q, Huang ZZ, Xu G, et al. Progress in the application of artificial intelligence in traditional Chinese medicine research. *Chin Tradit Pat Med*. 2024;46(10): 3529-3532. https://doi.org/10.3969/j.issn.1001 - 1528.2024.10.058.

185 Zhang WJ, Huai Y, Miao ZP, et al. Systems pharmacology for investigation of the mechanisms of action of traditional Chinese medicine in drug discovery. *Front Pharmacol*. 2019;10: 743. https://doi.org/10.1016/j.drudis.2006.11.008.

186 Wang SM, Long SQ, Wu WY. Application of traditional Chinese medicines as personalized therapy in human cancers. *Am J Chin Med*. 2018;46(5): 953-970. https://doi.org/10.1142/S0192415X18500507.

187 Zhang QR, Kong XJ, Xu HY, et al. Progress of studies on traditional Chinese medicine based on complex network analysis. *World J Trad Chin Med*. 2017;3(3): 28-37. https://doi.org/10.15806/j.issn.2311-8571.2016.0045

188 Hu YF, Wang ZL, Ni K, et al. Challenges in traditional Chinese medicine clinical trials: how to balance personalized treatment and standardized research? *Therap Clin Risk Manag*. 2025;21:1085-1094. https://doi.org/10.2147/TCRM.S523279.

189 Li DN, Hu J, Zhang L, et al. Deep learning and machine intelligence: new computational modeling techniques for discovery of the combination rules and pharmacodynamic characteristics of traditional Chinese medicine. *Eur J Pharmacol*. 2022;933:175260. https://doi.org/10.1016/j.ejphar.2022.175260.

190 Qu XL, Tian ZW, Cui JM, et al. A review of knowledge graph in traditional Chinese medicine: analysis, construction, application and prospects. *CMC-Comput Mat Contin*. 2024;81(3):3583-3616. https://doi.org/10.32604/cmc.2024.055671.

191 Wang X, Wang ZY, Zheng JH, et al. TCM network pharmacology: a new trend towards combining computational, experimental and clinical approaches. *Chin J Nat Med*. 2021;19(1):1-11. https://doi.org/10.1016/S1875-5364(21)60001-8.