Chinese Journal of Natural Medicines

Applications of artificial intelligence in the research of molecular mechanisms of traditional Chinese medicine formulas

Hongyu Chen, Ruotian Tang, Mei Hong, Jing Zhao, Dong Lu, Xin Luan, Guangyong Zheng, Weidong Zhang

Citation: Hongyu Chen, Ruotian Tang, Mei Hong, Jing Zhao, Dong Lu, Xin Luan, Guangyong Zheng, Weidong Zhang, Applications of artificial intelligence in the research of molecular mechanisms of traditional Chinese medicine formulas, Chinese Journal of Natural Medicines, 2025, 23(11), 1329–1341. doi: 10.1016/S1875-5364(25)60903-4.

View online: https://doi.org/10.1016/S1875-5364(25)60903-4

Chinese Journal of Natural Medicines 23 (2025) 1329-1341

Contents lists available at ScienceDirect

Chinese Journal of Natural Medicines

journal homepage: www.cjnmcpu.com/

Review

Applications of artificial intelligence in the research of molecular mechanisms of traditional Chinese medicine formulas

Hongyu Chen[a], Ruotian Tang[a], Mei Hong[a], Jing Zhao[a], Dong Lu[a], Xin Luan[a,*], Guangyong Zheng[a,*], Weidong Zhang[a,b,c,*]

[a] *Shanghai Frontiers Science Center for Chinese Medicine Chemical Biology, Institute of Interdisciplinary Integrative Medicine Research, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China*

[b] *Department of Phytochemistry, School of Pharmacy, Second Military Medical University, Shanghai 200433, China*

[c] *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China*

ARTICLE INFO

Article history:
Received 3 February 2025
Revised 24 March 2025
Accepted 21 May 2025
Available online 20 November 2025

Keywords:
Artificial intelligence
Traditional Chinese Medicine Formula
Molecular Mechanism
Machine learning
Knowledge graph

ABSTRACT

Traditional Chinese medicine formula (TCMF) represents a fundamental component of Chinese medical practice, incorporating medical knowledge and practices from both Han Chinese and various ethnic minorities, while providing comprehensive insights into health and disease. The foundation of TCMF lies in its holistic approach, manifested through herbal compatibility theory, which has emerged from extensive clinical experience and evolved into a highly refined knowledge system. Within this framework, Chinese herbal medicines exhibit intricated characteristics, including multi-component interactions, diverse target sites, and varied biological pathways. These complexities pose significant challenges for understanding their molecular mechanisms. Contemporary advances in artificial intelligence (AI) are reshaping research in traditional Chinese medicine (TCM), offering immense potential to transform our understanding of the molecular mechanisms underlying TCMFs. This review explores the application of AI in uncovering these mechanisms, highlighting its role in compound absorption, distribution, metabolism, and excretion (ADME) prediction, molecular target identification, compound and target synergy recognition, pharmacological mechanisms exploration, and herbal formula optimization. Furthermore, the review discusses the challenges and opportunities in AI-assisted research on TCMF molecular mechanisms, promoting the modernization and globalization of TCM.

1. Introduction

Traditional Chinese medicine (TCM) embodies millennia of medical knowledge and maintains a crucial role in preserving the health of the Chinese population throughout history. Traditional Chinese medicine formulas (TCMFs), characterized by their multi-component, multi-target, and multi-pathway properties, demonstrate particular effectiveness in treating complex and chronic diseases [1]. A notable example is artemisinin, a sesquiterpene lactone extracted from Artemisia annua, which has exhibited significant anti-malarial properties and saved numerous lives globally [2]. During the COVID-19 pandemic, TCMFs play a significant role in patient care. A key aspect of this contribution is the implementation of the “three medicines and three formulas” (San Yao San Fang). These interventions effectively alleviated clinical symptoms and accelerated recovery among SARS patients [3, 4]. In modernizing TCMFs, researchers have proposed various strategies to understand the holistic regulatory mechanisms underlying their therapeutic efficacy. Academics Boli Zhang and Yongyan Wang advocate a component compatibility-based strategy for TCM formulation [5]. This approach emphasizes “identifying primary components, enhancing principal therapeutic effects, and reducing adverse effects” to develop new formula compositions guided by component compatibility. Professor Weidong Zhang’s research team [6] presents a system-oriented strategy for TCMF pharmacology investigation, aiming to analyze pharmacological effects through integrated modern analytical technologies, including chemical substance profiling, compound pharmacokinetic (PK) assay, and component PK-pharmacodynamic (PD) correlation analysis. Dean Guo’s team [7] establishes a systematic analytical framework for active component identification in TCMFs, utilizing spatial metabolomics based on mass spectrometry imaging to comprehensively analyze component distribution and tissue-specific metabolic patterns *in vivo*. Nevertheless, understanding the pharmacological mechanisms remains challenging due to the multi-component, multi-target, and multi-pathway nature of TCMFs.

Since the 1990s, artificial intelligence (AI) techniques, including natural language processing (NLP) [8-13], machine learning (ML) [14-18], computer vision (CV) [19-24], and robotics [25-29], have achieved significant breakthroughs. AI technologies demonstrate substantial impact in the biopharmaceutical sector, influencing drug discovery, efficacy monitoring, pharmacological analysis, and health promotion [30-34]. For instance, ML models incorporating na-

---

* Corresponding author.
E-mail addresses: luanxin@shutcm.edu.cn (X. Luan); gyzheng@shutcm.edu.cn (G. Zheng); wdzhangy@hotmail.com (W. Zhang)

https://doi.org/10.1016/S1875-5364(25)60903-4

ive bayes (NB), decision tree, logistic regression, and support vector machine (SVM) algorithms have been employed to identify anti-psychotic DDIs [35]. ML and knowledge graph (KG) technologies facilitate new indication discovery for approved drugs by integrating various information, including structural characteristics, affected pathways, and molecular mechanisms [36, 37]. Deep learning (DL) demonstrates particular advantages in drug development, enhancing discovery efficiency, cost effectiveness, and success rates in identifying potential drug candidates [38-40]. Exscientia developed a PKC theta inhibitor for immunology and inflammation diseases within 11 months using DL technology [41]. Insilico Medicine identified a novel inhibitor (ISM001-055) targeting anti-fibrotic pathways in just 9 months using their DL-driven drug discovery pipeline [42]. This acceleration in drug R&D stems from AI's efficient handling of multi-dimensional data, enabling probabilistic modeling of polypharmacological interactions.

AI applications in TCM modernization encompass the four diagnostic methods, syndrome differentiation, and digital preservation of TCM knowledge (Table 1). AI advances have enhanced the Chinese medicine industry's intelligent processes, particularly in ingredient identification, component validation, herb cultivation, quality standardization, auxiliary diagnosis, health management, target prediction, mechanism dissection, clinical efficacy evaluation, and personalized formulation (Fig.1). DL has accelerated multi-omics data analysis and scientific law discovery, enabling precise therapeutic effect assessment for TCMFs [43, 44]. However, TCMFs' holistic nature and complex component-target interactions present significant challenges in elucidating their molecular mechanisms. AI shows promise in addressing these challenges through analyzing ingredient combinations, interpreting component-target interactions, and revealing multi-target synergistic effects [45-47].

This review explores the applications of AI in studying molecular mechanisms of TCMFs. Key area of focus include compound ADME prediction, molecular target identification, synergy effect elucidation, pharmacological mechanism interpretation, and formula component optimization. The discussion delves into the challenges and future prospects of AI-assisted mechanistic investigations of TCMFs. By integrating ancient wisdom (TCM) with modern technology (AI), this approach has the potential to significantly enhance human healthcare.

2. Applications of AI in the study of molecular mechanisms of TCMFs

2.1. ADME prediction of components for TCMFs

The PK behavior of drugs in the human body can be characterized by key parameters related to ADME, such as solubility, permeability, plasma protein binding (PPB), oral bioavailability, blood-brain barrier (BBB) permeability, human intestinal absorption (HIA), area under the plasma concentration-time curve (AUC), urinary excretion, and half-life. These parameters are essential in determining the viability of candidate drug development [55]. In the context of TCM formulas, ADME and PK studies serve a vital role in elucidating their PD material basis and underlying therapeutic mechanisms [56-58]. However, the spatial conformation complexity and physicochemical property diversity of natural products (NPs) and formula components present significant challenges for the systematic analysis of their PK characteristics.

In recent years, AI-driven methods have demonstrated significant potential in addressing these challenges, facilitating efficient identification of TCM components. For example, Zhang et al. developed a BBB permeability prediction model through ML algorithms, specifically designed for NPs derived from TCM [59]. This model addressed the limitations of conventional models based on synthetic drugs in TCM small molecules prediction. By combining the SVM and random forest (RF) algorithms, the model achieved an accuracy of 81% in predicting BBB permeability for TCM-derived compounds. Furthermore, the model utilized key factors affecting BBB permeability (such as lipophilicity, hydrogen bond count, and molecular polarity) to accurately predict differences in permeability between predecessor and descendant metabolites. Guo et al. employed the admetSAR platform to conduct virtual ADME property screening on 159 compounds identified from Xiao'er Fu Pi Granules (XEFP), encompassing flavonoids, glycosides, alkaloids, and other components [60]. The analysis identified 129 out of 159 compounds with favorable absorption and low toxicity profiles. Wang et al. developed TCM-ADMEpred, a novel multi-component PK prediction model for TCM formulas [61]. By integrating PK data and the structure information of TCM components, the model employed the trapezoidal rule to calculate the AUC time curves of seven components from Yuanhu Zhitong Formula (YZP). Additionally, the introduction of Γ + was used to quantify PK correlations among structurally similar components, enabling the successful prediction of the metabolic behavior of three alkaloids in YZP. This study demonstrates the potential of incorporating structural information into PK modeling to provide quantitative insights into the complex pharmacokinetics of TCMFs.

While traditional ADME prediction methods, such as PBPK models, can simulate compound absorption, distribution, and metabolism, their application to TCM systems remains limited due to sophisticated components and intricate pharmacokinetics of these systems. For instance, the absorption and transport (CAT) model within the PBPK framework can simulate drug absorption across different anatomical regions of the small intest-

Table 1 Applications of AI in TCM.

Application	Algorithm	Description	Reference
Tongue diagnosis	Faster R-CNN	A model is constructed using the Faster R-CNN algorithm, which infers the clinical syndromes of subjects by learning the local features of teeth-marked tongues and cracked tongues from tongue image data.	Liu et al., 2019 [48]
Tongue diagnosis	GoogleNet, ResNet	DL modeling is employed to classify and assess three primary tongue features in clinical patients: teeth marks, cracks, and tongue coating thickness.	Song et al., 2021 [49]
Face diagnosis	DL	An automatic TCM facial complexion diagnosis algorithm based on multi-feature fusion of facial characteristics, addressing the issue of TCM facial complexion classification.	Lin et al., 2021 [50]
Eye diagnosis	GAN	A retinal vessel segmentation method combining with the generative adversarial network (GAN) is proposed for the classification and diagnosis of patient eye images.	Li et al. ,2020 [51]
Syndrome Classification	KNN, FCNN, and 1D-CNN	Three AI algorithms (i.e., multi-label k-nearest neighbors, fully connected neural networks, and one-dimensional convolutional neural networks)are utilized to explore 28 common clinical syndromes of TCM zang-fu differentiation.	Du et al., 2023 [52]
Knowledge integration	KG	A large-scale KG is built to reflect characteristics of Yin Deficiency Syndrome through integrating medical texts and physicians, principles and methods, prescriptions and medicinals, as well as syndrome differentiation and treatment.	Zeng et al., 2023 [53]
Knowledge integration	KG	A KG about the Shanghan Lun is designed and constructed, utilizing TCM named entity recognition (NER) technology to mine literature related to Yangming disease.	Kuang , 2022 [54]

Note, CNN: Convolutional Neural Network; KNN: K-nearest neighbor; BP: Back Propagation; GAN: Generative Adversarial Network; KG: Knowledge Graph; FCNN: fully connected neural network.

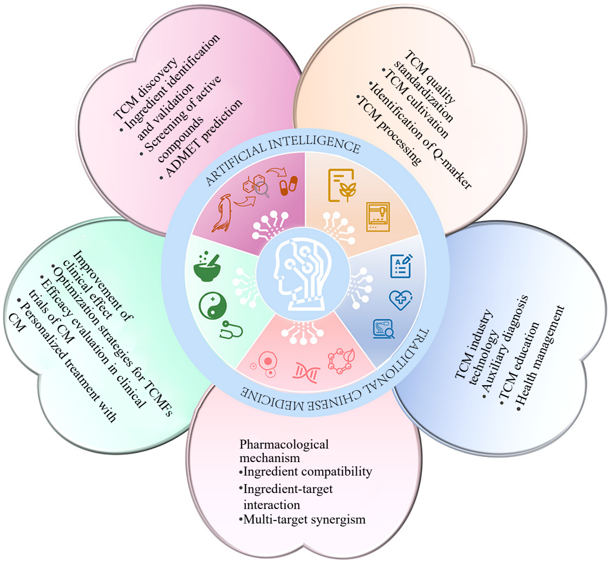

Fig. 1 Applications of AI in the TCM industry.

ine but lacks accuracy in accounting for dynamic factors, such as dissolution rate, pH-dependent solubility changes, first-pass effects, and drug degradation. These limitations result in reduced predictive accuracy for drugs with low solubility or permeability [62]. In contrast, AI technologies, particularly graph neural networks (GNNs) [63, 64], excel in capturing dynamic molecular behaviors of compounds, such as physicochemical properties, molecular representations, and drug-like characteristics. For instance, Yang et al. demonstrated that GCNN models surpassed traditional RF models in PD studies (e.g., rat PPB, human PXR activation), achieving a 14% improvement in R^2 values [65]. Similarly, Liu et al. applied multi-task GCNN models to solubility and PXR-related investigations, consistently outperforming Cubist models across various metrics [66]. Comparisons between traditional and AI-based compound ADME prediction are summarized in Table 2 [67-69]. AI-based ADME prediction approaches are more appropriate for PD study of TCM components since these approaches are superior to traditional methods in data processing, mathematical modeling, and dynamic prediction. Currently, several online platforms that integrate AI technologies have been developed for ADME/PK research for TCM (Table 3) [70-77]. Although the reliability of ADME models in identifying the PD material basis of TCMFs remains a subject of debate, particularly regarding their ability to capture interactions among complex components, such platforms have significantly accelerated PK studies of TCMFs. AI models predict the dynamic behavior of TCM components through incorporating molecular and conformational information (e.g., molecular interaction forces, conformational energy), thereby enhancing model generalization and predictive accuracy. This approach accelerates investigations into the PD material basis of TCMFs and facilitates the elucidation of their mechanisms of action, ultimately supporting their clinical translation and application.

2.2. Molecular target prediction for TCMFs

Current ADME prediction models facilitate the efficient identification of compounds with favorable PK profiles, forming a critical foundation for constructing component–target interaction networks and dissecting the molecular mechanisms underlying TCMFs. Fully elucidating the "multi-component, multi-target" nature of TCMFs requires explicit characterization of the interactions between bioactive compounds and their corresponding molecular targets [78-80]. Traditionally, molecular targets of TCM compounds are identified through biological experiments, which is a time-consuming and costly process. Moreover, conventional experimental methods often encounter difficulties in systematically identifying the multiple molecular targets associated with the various active compounds of formulas, limiting understanding of the target clusters and their synergistic interactions.

Biological processes induced by TCMFs can be modeled *via* heterogeneous molecule networks, facilitating network-driven identification of therapeutic targets for these formulas. During the network construction process, TCMFs-related disease symptoms, compatibility herbs, active components, molecule targets, and multi-omics information are retrieved from public databases (Table 4). In detailed investigations, Qi et al. utilized structure-

Table 2 Comparisons between traditional and AI-based methods for compound ADME prediction.

Metric	Traditional methods	AI-based methods	Reference
Data Processing	Relies on manually curated datasets and handcrafted feature selection	Employs automated data preprocessing and feature extraction	67-69
Mathematical Modeling	Captures low-level feature interactions using linear or basic non-linear regression approaches	Learns complex, high-order non-linear relationships via advanced neural networks	
Dynamic Prediction	Restricted to static, single-timepoint measurements	Enables dynamic, time-resolved PK simulations	

Table 3 Platforms for compound ADME property investigation.

Name	Description	Website	Reference
DrugBank	A comprehensive database of drugs, presenting bioinformatic and cheminformatic data of compounds and their targets.	https://go.drugbank.com/	Wishart et al., 2006 [70]
TCMSP	A unique systemic pharmacology platform of Chinese herbal medicines that captures the relationships between drugs, targets, and diseases.	https://tcmsp-e.com/	Ru et al., 2014 [71]
PkCSM	A ML-based online tool for PK and PD property prediction for compounds.	https://biosig.lab.uq.edu.au/pkcsm/	Pires et al., 2015 [72]
CancerHSP	A database contains 2439 anticancer herbal medicines with 3575 tumor suppressive ingredients. For each ingredient, the molecular structure and nine key ADME parameters are provided.	https://lsp.nwu.edu.cn/rj1/CancerHSP2.htm	Tao et al., 2015 [73]
Swiss ADME	An ADME prediction toolkit comprising a suite of computational models for assessing key pharmacokinetic and safety-related properties of compounds, including solubility, plasma protein binding, hepatic metabolism, renal excretion, CYP450 interactions, toxicity, and overall safety profiles.	http://www.swissadme.ch/	Daina et al., 2017 [74]
ADMETlab 2.0	A platform for systematical evaluation of ADMET properties, as well as physicochemical and medical characteristics for compounds.	https://admetmesh.scbdd.com/	Xiong et al., 2021 [75]
TCMIP	An integrative pharmacological research platform of TCM, including standardized information on the widely used herbs and formulas of TCM.	http://www.tcmip.cn/TCMIP/index.php	Wang et al., 2021 [76]
ADMETlab 3.0	An updated web-based platform for evaluating ADMET properties, physicochemical parameters, and medicinal chemistry features, offering improved accuracy, expanded data coverage, API access, and uncertainty estimation to streamline drug discovery.	https://admetlab3.scbdd.com	Fu et al., 2024 [77]

similarity prediction methods to construct interaction networks between active compounds and molecular targets of Xuefu Zhuyu Decoction (XFZYD) and Gualou Xiebai Banxia Decoction (GLXBD) [81]. They subsequently collected disease target sets of coronary heart disease (CHD) and identified the top ten molecular targets for XFZYD and GLXBD in CHD treatment through network topology analysis. Liu et al. employed network pharmacology to investigate the molecular mechanisms and therapeutic targets of Huangqin Decoction (HQD) in treating ulcerative colitis [82]. Through combining HQD's active compound data, gene differential expression data, and protein-protein interaction (PPI) data, they identified interleukin-6 (IL-6), tumor necrosis factor (TNF), IL-1B, PTGS2, ESR1, and PPARG as core targets, revealing key signaling pathways (TNF signaling, IL-17 signaling, and Th17 cell differentiation) in ulcerative colitis treatment. Huali et al. constructed a target-pathway network for Yupingfeng Decoction (YPFD) and employed centrality metrics to evaluate target importance, ultimately selecting the top 10% key targets to demonstrate the formula's pharmacological mechanisms. Liu et al. developed an ingredient-target network for Danggui Niantong Decoction (DGNTD) and expanded it into a "herb-compound-pathway-target" network, identifying main active compounds (quercetin, kaempferol, paeoniflorin, wogonin, and nootkatone) and core targets (CASP8, CXCL8, FOS, IL-1B, IL-6, JUN, PTGS2, STAT1, and TN) for the formulas.

2.3. Synergistic effect revelation for TCMFs

Following molecular target recognition for active components, elucidating component and target coordinated interactions and deciphering their collective therapeutic effects are essential for the mechanistic exploration of TCMFs. Modern medical research shows that combination therapies are more effective than single-drug treatments. This is because they can regulate multiple key aspects of diseases, including various pathogenic targets, compensatory signaling pathways, and complex molecular networks [94]. This approach reduces drug resistance and side effect risks while enabling lower drug doses to achieve equivalent or superior therapeutic effects (Fig. 2) [95-96]. Synergistic effects primarily fall into two categories: drug combination and target synergy [97, 98]. Drug combination represents a therapeutic strategy where multiple drugs interact synergistically, producing effects exceeding their expected additive impact [99, 100]. Target synergy occurs when multiple drug targets are simultaneously activated or inhibited, resulting in a combined therapeutic effect that surpasses the sum of their individual actions [101]. Synergistic effects align closely with TCM's "Qi Qing He He" (Harmony of Seven Emotions) compatibility theory, where multiple formula components work together through multi-target and multi-pathway synergies to produce therapeutic effects. Compared to Western medicine, drug synergy in TCMFs encompasses broader implications. TCM synergy involves both micro-level active compound interactions and macro-level medicinal herb synergies. Additionally, the extensive target range affected by TCMFs' multiple components significantly increases the complexity of understanding synergistic mechanisms within their target networks. Consequently, innovative research strategies are necessary to better elucidate TCMFs' overall synergistic mechanisms.

With rapid development of AI, network pharmacology and ML techniques [40, 102] have increasingly been applied to investigate TCMFs' complex synergistic effects (Fig. 3). These methods provide a comprehensive framework for exploring TCM molecular mechanisms and may help unravel the mysteries of "Yi Yin Tang Ye" (TCMFs). Current studies typically begin by constructing network models that connect drugs, targets, and diseases using public data. In these models, core drugs, key target nodes, and their interactions are analyzed using network topological metrics such as node degree, betweenness centrality, and clustering coefficient to reveal drug synergies [103, 104]. Wang et al. implemented a network pharmacology framework to explore herbal compatibility and identify synergistic compound interactions through network topology analysis [105]. By integrating herb, compound, and target relationships, they developed herb-herb distance metrics (nearest, shortest, and central distances) to evaluate herb-ingredient and ingredient-target interactions in binary networks, investigating herb pairs' molecular mechanisms. Gan et al. mapped disease-specific and TCM-associated targets into the human PPI network to examine synergistic effects, clarifying

Table 4 Commonly used databases in the pharmacological study of TCMFs.

Database Name	Description	Data Source	Reference
TCMID	A comprehensive database providing correlated information between Traditional Chinese Medicine and modern life sciences.	Text-mining TCM-ID; HIT; TCM@Taiwan; HIT; STITCH; OMIM; DrugBank; OMIM; STITCH;	Wishart et al., 2013 [83]
TCMSP	A system pharmacology platform presenting herbal chemical constituents, ADME properties, and multi-dimensional drug-target-disease network information of TCM.	Chinese pharmacopoeia; PubChem; ChEMBL; ChemSpider; DrugBank TTD; PharmGKB; HIT	Ru et al., 2014 [71]
BATMAN-TCM	An integrative platform about TCM pharmacological mechanisms	TCMID; TCM-Suite; LTM-TCM; ITCM; TCMID; TCM-Suite; HERB; ITCM TCMID; TCM-Suite; HERB; ITCM; SWISS-PROT; KEGG; Gene Ontology; OMIM/TTD DrugBank; KEGG; TTD; HIT; HERB	Liu et al., 2016 [84]
ETCM	A system pharmacology database offering standardized clinical profiles, multi-omics results, and quantitative ingredient data.	Text-mining; Pubchem; HPO; OMIM; DisGeNET; ORPHANET; Reactome; HPRD; MINT; IntAct Molecular Interaction Database; DIP	Xu et al., 2019 [85]
SymMap	A TCM database providing massive information on herbs, ingredients, targets, as well as clinical symptoms and diseases.	Chinese pharmacopoeia; TCMID; TCMSP; HIT; HPO; DrugBank; OMIM; MeSH; SIDER; UMLS; Text-mining	Wu et al., 2019 [86]
HERB	A high-throughput experiment- and reference-guided database of traditional Chinese medicine.	SymMap; TCMSP; TCM-ID; SciFinder; PubChem; text mining; GeneCards; TTD; DisGeNet; HPO; Disease Ontology databases; HIT	Fang et al., 2021 [87]
ITCM	A bioinformatics platform combining multi-modal TCM research data, featuring literature mining, small-molecule expression profiling, and computational toolkits.	SYMMAP; TCMID; TCMSP; ETCM; NPASS; CMAUP; TCMIO; HERB; DisGenet; DrugBank; TCGA; DGIdb; CTD	Tian et al., 2021 [88]
DCABM-TCM	A database collecting blood constituents of TCM prescriptions and herbs.	Text-mining; TCMID; Entrez Gene databas; KEGG; OMIM; CTD; Swiss-Prot; BATMAN-TCM	Liu et al., 2021 [89]
TCMBank	A standard TCM data platform encompassing herbs, active ingredients, disease associations, and gene targets.	TCMID; TCMSP; SymMap; HERB; ETCM; text-mining	Lv et al., 2023 [90]
ccTCM	A quantitative platform presenting the distribution, similarity, and molecular mechanism of components and compounds	the Chinese Pharmacopoeia; PubChem; ChEMBL; ZINC	Yang et al., 2023 [91]
IGTCM	A multi-omics platform bridging TCM and molecular biology through herb-genome-constituent networks	NCBI; GPGD; SymMap; HERB	Ye et al., 2023 [92]
TCMM	A TCM database encompassing prescription, ingredient, target, and biological relation information	PrimeKG; PharMeBINet; TCMBank; SymMap; CPMCP; TCMID	Ren et al., 2024 [93]

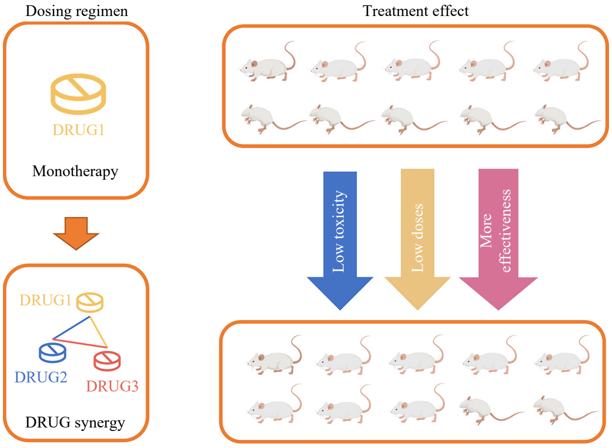

Fig. 2 Advantages of drug combination therapy.

system-level mechanisms underlying herb-symptom interactions [106]. Syn-COM represents a novel multi-layer collaborative framework integrating network pharmacology and AI inference for TCM investigation [107]. This framework screens effective ingredient combinations and synergistic effects of herbs for treating gouty arthritis (GA), utilizing Si and Ec indices to represent the "Emperor, Minister, Assistant, and Envoy" principles of herb compatibility, exploring active compounds' primary and secondary roles in TCMFs.

Deep Learning (DL), with its capabilities in feature extraction, automatic modeling, and correlation prediction, provides an effective approach for understanding the complex interactions between multiple drugs and targets [44, 101, 108]. In synergy studies of TCMFs, DL models help identify key compounds and targets that contribute to synergistic effects. While traditional TCM theories interpret specific herb pairs in TCMFs based on their medicinal properties in clinical applications, the molecular mechanisms underlying these synergistic effects remain in early research stages [109]. Xue et al. implemented DL methods to examine herb pairs, incorporating decoction properties, taste, and meridian attributes into the dataset. The knowledge-enhanced semantic representation (ERNIE) model analyzed the contextual information on herb pairs [110]. The ERNIE model demonstrated superior accuracy in identifying potential herb pairs from TCM datasets compared to CNN, RNN, and BERT models. Yang et al. developed an association graph connecting TCM herbs, compounds, and proteins for SARS-CoV-2 treatment [111], establishing a heterogeneous "herb-compound-protein" (HCP) network through meta-path embedding. The HCP network identified candidate herbs targeting structural, non-structural, and accessory proteins of SARS-CoV-2. Subsequently, the variational graph autoencoder (VGAE) model recommended 20 herb pairs as potential COVID-19 treatments.

Most existing computational models focus on predicting synergistic effects between two synthetic drugs or compounds, limiting their effectiveness when applied to multiple components of TCMFs. To overcome this limitation, researchers have developed new approaches integrating diverse data types (including chemical structures, biological data, and textual information) to

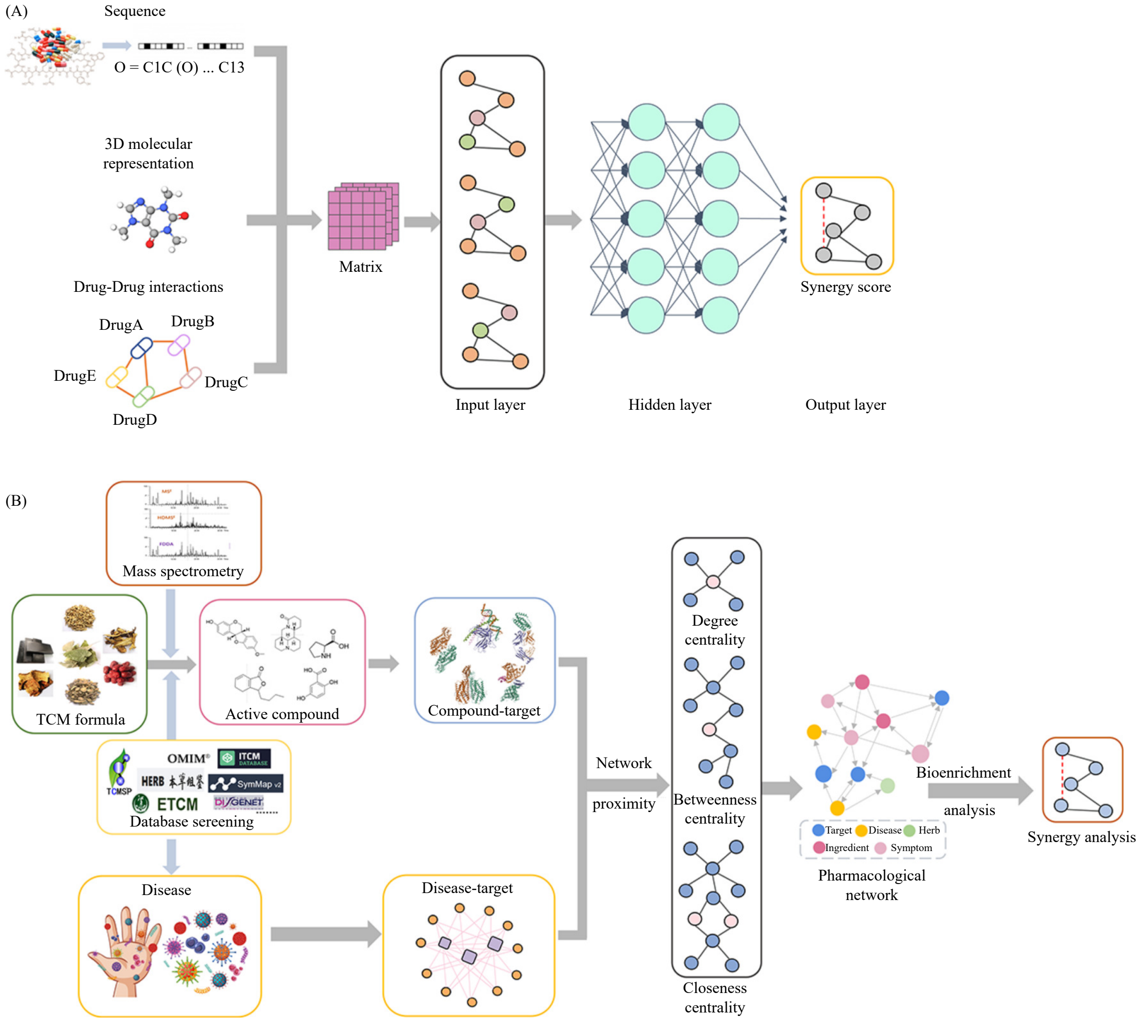

Fig. 3 Strategy of synergistic effect investigation in TCMFs based on DL algorithm (A) and network pharmacology approach (B).

comprehensively evaluate multi-component combinations in TCM [112-115]. Guo and colleagues introduced a novel framework to demonstrate the synergistic effects of multiple TCMF components [60]. Their approach utilized unsupervised ML algorithms to establish connections between compound modules and cellular functions based on FDA-approved drugs, identifying active components in Xiao'er Fupi Granules. They extracted "compound-target-cell-drug" meta-paths from compound modules with similar functions, correlating these modules with cellular functions to elucidate synergistic effects. Sun and colleagues developed a heterogeneous network method for analyzing compound combination mechanisms, drawing inspiration from complex networks and overlapping community detection [116]. They employed the Doc2vec model to represent compound semantic properties and utilized the Jaccard index to measure compound attribute similarities, providing insights into synergistic patterns between traditional Chinese and modern medicines.

AI methodologies promote synergistic effect studies between components and targets, particularly in predicting specific interactions. These correlated interactions serve as intermediate information for mechanism exploration of TCMFs.

2.4. Pharmacological mechanism exploration of TCMFs

Network models have been widely used to uncover the pharmacological mechanisms of traditional Chinese medicine formulations (TCMFs). By leveraging holistic organism-level data obtained through multi-omics technologies during TCM interventions, these models provide valuable insights into how TCMFs work [117, 118]. Xiong et al. developed an integrated strategy combining serum chemistry, metabolomics, and network pharmacology to identify the material basis and mechanisms of Danggui Shaoyao San (DSS) in treating premenstrual dysphoria (PD) [119]. Lu et al. combined network pharmacology and transcriptomics to elucidate the active ingredients, targets, and potential therapeutic mechanisms of Ningfei Pingxue Decoction (NFPX) for acute respiratory distress syndrome (ARDS) treatment [120]. Jiang et al. investigated Shentong Zhuyu Decoction (STZYD) mechanisms against rheumatoid arthritis (RA) using metabolomics and network pharmacology, identifying 10 key metabolites, targets and pathways [121]. Ma et al. utilized a multi-omics integration approach (incorporating transcriptomics, proteomics, and metabolomics) to reveal the molecular mechanisms of Qijiao Shengbai Capsules (QSC) in treating leukopenia [122]. Moreover, intelligent network models have played a significant role in advancing the prevention and treatment of COVID-19 and provided valuable insights into understanding the disease [123, 124]. Zhao et al. developed an integrated biological model to explain the molecular mechanisms of Qingfei Paidu Decoction (QFPDD) in treating COVID-19 by incorporating data on blood-migrating components, metabolic pathways, target organs, and clinical symptoms [125].

To overcome the limitations of traditional network models in analyzing node and edge properties without enabling generalized predictions, Qi et al. developed a GNN-based framework that employs multi-label classification and link prediction models to explore the pharmacological basis and mechanisms of Xuefu Zhuyu Decoction (XFZYD) [126]. In practice, they conceptualized the material basis of XFZYD in treating both stable and unstable angina as a multi-label classification problem involving GNN nodes and framed the action mechanisms as a link prediction task between edges. They constructed a correlative network representing compound, target, and pathway relationships through node similarity calculations and applied the Node2Vec algorithm for network embedding. Subsequently, they employed the multi-label classification model to explore potential therapeutic mechanisms of XFZYD. DeepDGC, a DL-based network pharmacology framework incorporating ADMET predictions, molecular docking, molecular dynamics simulations, and compound and target structure modeling, was developed to investigate licorice's molecular mechanism in treating COVID-19 [127]. Zhang and colleagues performed a systematic analysis of Danshen and Chuanxiong's chemical composition in Guanxinning Tablets, identifying more than 40 major chemical components [128]. Using AI-based network pharmacology methods, they predicted these components' potential targets and investigated the associated signaling pathways. They found that Danshen and Chuanxiong's chemical components primarily affect VEGF and HMGB1 pathways, regulating glucose and lipid metabolism, inhibiting myocardial fibrosis, preventing thrombosis, and reducing inflammation and cell necrosis. Furthermore, their examination of multi-scale biological effects revealed that salvianolic acid inhibits COX-1 expression, reduces platelet aggregation, and limits myocardial infarction extent, while ligustilide blocks the coagulation cascade, inhibits thrombosis, and works synergistically with tanshinone to enhance blood circulation [128].

Network pharmacology typically integrates multi-omics data (e.g., transcriptomics, proteomics, metabolomics) to elucidate molecular mechanisms through which active compounds and their targets produce therapeutic effects. However, multi-omics data integration presents several challenges, including data incompleteness, complexity, and model opacity. Data incompleteness, often resulting from poor sample quality and insufficient sample sizes, can be addressed through AI-based data imputation methods. For example, Eltager et al. developed a K-nearest neighbors (KNN) variant algorithm that utilizes single-cell RNA sequencing (scRNA-seq) data to impute missing single-cell assays for transposase-accessible chromatin (scATAC-seq) profiles [129]. Data complexity can lead to the "curse of dimensionality" when integrating various omics data [130]. To address this challenge, researchers have developed AI-based graph linking strategies to reduce data dimensionality and incorporate features from various omics [131]. Regarding model opacity, KG technology has been introduced to TCM research, as it enhances model interpretability when combining omics data. A KG represents a semantic network revealing entity relationships, depicting concepts and their connections in symbolic form. Its fundamental unit consists of an "entity-relation-entity" triple, where relationships link entities to form a knowledge web [132-134]. TCM-KGs consolidate extensive information from classical texts and modern literature, encompassing TCM theories, terms, syndromes, diseases, prescriptions, herbs, compatibility, usage, and dosage. These KGs address the challenge of fragmented knowledge and facilitate various information integration in TCM, thereby enhancing TCM knowledge accessibility and applicability [135-137]. KG construction involves extracting valuable data from heterogeneous sources and integrating them into a unified semantic network, encompassing data acquisition, information extraction, and knowledge fusion and representation. For instance, TarKG was developed through the integration of 7 biomedical KGs, 9 public databases, and 11 TCM knowledge repositories, incorporating multiple TCM-specific relationships, including herb-compound, TCM symptom-symptom, herb-gene, herb-disease, TCM formula-disease, TCM symptom-disease, and TCM syndrome-disease associations [138]. These relationships across TCM entities provide substantial information for exploring TCMF mechanisms. Additionally, Zhu et al. developed HerbKG, focusing on four key entities: herbs, compounds, genes, and diseases [139]. This graph addresses gaps in existing TCM-KGs, which primarily focus on diagnostic and treatment knowledge while often overlooking molecular-level information. Using a deep transfer learning framework, HerbKG integrated 53,754 TCM-related triples, providing a comprehensive knowledge network for molecular research in TCM. Notably, recent DL advances have enabled TCM KG construction without manually annotated features. Named entity recognition

(NER) models, incorporating various TCM resources (such as books, articles, and records), employ self-attention mechanisms to automatically identify TCM entities, significantly enhancing KG construction efficiency and accuracy [140].

2.5. Optimization of composition for TCMFs

In the theoretical framework of TCM, the concept of *Qi Qing He He* (Harmony of Seven Emotions) encompasses both individual herb applications and compatibility principles between medicinal materials, reflecting TCM's holistic and dialectical philosophy while aligning with modern medicine's approach to synergistic drug combinations for toxicity reduction and efficacy enhancement [103, 104]. Unlike Western medicine's straightforward drug combinations, TCMFs are formulated based on extensive clinical experience and compatibility principles, wherein multiple herbs function synergistically to treat diseases through syndrome differentiation. The inherent complexity of multi-component, multi-target formulas present significant challenges in identifying and validating herbal combination efficacy and safety. The component-based Chinese medicine research paradigm, introduced by Academics Boli Zhang and Yongyan Wang, marks a crucial shift from traditional macro-level TCMF optimization to precise micro-level component quantification. This approach establishes a solid theoretical foundation for modernized TCMF optimization [5, 80]. Following the paradigm's "two well-defined aspects" principle (well-defined material basis and mechanism of action), the material basis encompasses: (1) active components that provide the foundation for pharmacological effects, and (2) functional components that serve auxiliary roles such as solubilization, toxicity reduction, or efficacy enhancement. The well-defined mechanism of action requires modern scientific methods to elucidate relationships between active components and molecular targets, as well as component interactions, establishing scientific evidence for TCM efficacy. The paradigm suggests that formula optimization centers on designing optimal relationships between component dosage, ratio, and pharmacological efficacy, considering toxicity and side effect predictions. The application of advanced AI technologies in component-based TCMF optimization represents a crucial pathway for advancing beyond traditional empirical formulation approaches and innovating modern Chinese medicines, offering a breakthrough direction for new TCMF development [104-107].

TCM's theoretical framework emphasizes the principle of complementary drug use while avoiding antagonistic combinations, suggesting that when toxicity occurs, counterbalancing agents should be employed. This principle acknowledges that while adverse effects may occur in compound formula applications, they can be mitigated. Approximately 230 000 TCM-related adverse reaction cases are reported globally annually [141], emphasizing the necessity of identifying toxic components in herbal formulas to optimize prescriptions, minimize toxicity, and enhance efficacy. He et al. addressed this challenge by integrating 13 molecular fingerprint features and implementing eight ML algorithms to develop compound toxicity prediction models [142]. Their methodology incorporated recursive feature elimination (RFE) for optimal feature subset selection in model construction, successfully identifying hepatotoxic compounds in TCMFs through a herb-hepatotoxic compound network using the HILI dataset. Wu et al. developed an innovative multi-task deep neural network incorporating element-specific persistent homology (ESPH) for quantitative compound toxicity prediction. ESPH maintains essential chemical features during topological abstraction, addressing small molecule representation challenges [143]. Their findings demonstrated that combining compound topological and physical descriptors with advanced ML algorithms, such as DNN, RF, and gradient boosting decision trees, achieved superior performance across various tasks related to compound toxicity prediction. Liu et al. applied DL techniques to predict compound toxicity and side effect, with a particular focus on safety assessment in COVID-19 treatment [124]. In recent years, AI models such as DeepHit, FP-ADMET, and ResNet18DNN (a variant of deep neural networks) have successfully been implemented to predict various toxicological properties such as hERG inhibition, LD50, drug-induced liver injury (DILI), Ames mutagenicity, carcinogenicity, skin sensitization, and Tox21 assay endpoints [144].

A comparative analysis of representative AI models in compound toxicity prediction is shown in Table 5 highlighting their methodological innovations, performance benchmarks, and practical applications in TCM research. Collectively, AI models offer valuable insights into ensuring the safe application of Chinese medicine while also contributing to the optimization of TCMFs.

2.6. Intelligent recommendation of TCMFs

In the field of TCMF recommendations, FordNet, an advanced intelligent system, has garnered significant academic attention [148]. This system utilizes CNNs to extract high-level features from over 20 000 medical case records collected between 2013 and 2020, subsequently establishing a heterogeneous network of herbs, ingredients, and targets. Through NE techniques, FordNet identifies underlying patterns of TCMFs and recommends new formulas tailored to specific diseases. Compared to models utilizing only macro-level data, FordNet incorporates both macro- and micro-level information for active ingredients, archieving an 17.3% increasement of hit rate compared to the baseline method. To assess the clinical application potential of FordNet, formula recommendations were conducted for 50 gastrectasia patients based on electronic health records (EHRs). The results indicate that FordNet can effectively recommends appropriate formulas aligned with TCM expert experience. Beside, the MGCN recommendation model of TCMFs enhances predictive accuracy by employing multiple GNNs to model the relationships between symptoms, physiological states, and TCM syndrome patterns [149]. This model consists of two main modules: a TCM feature aggregation module and an herbal medicine prediction module. The TCM feature aggregation module establishes a symptom-state element-symptom graph (Se) and a symptom-syndrome-symptom graph (Ts), effectively modeling the complex multivariate relationships between symptoms and prescriptions. The herbal medicine prediction module, utilizing state elements, syndrome types, and symptom data, employs a multi-layer perceptron (MLP) to generate formula recommendations. Despite these advances, AI-driven TCM prescription systems encounter several critical challenges. First, data quality remains a significant concern, as incomplete or inaccurate data may result in incorrect recommendations. Moreover, access to high-quality clinical datasets in TCM remains limited. Another major challenge is model interpretability—non-transparent AI models impede user understanding and diminish trust. Additionally, clinical validation of AI-recommended TCM prescriptions is essential for ensuring safety and efficacy. To address these challenges, researchers increasingly focus on enhancing the transparency and accuracy of AI models in TCM prescription recommendations. Jin et al. introduced the Metapath-guided graph attention network (MGAT) to improve the interpretability of herbal medicine recommendations [150]. The model contributes to their broader initiative of transitioning TCM from an empirical medicine framework to an evidence-based system, integrating modern pharmacological findings with traditional TCM knowledge into a comprehensive KG. MGAT employs predefined KG metapaths to guide network neighbor selection during information propagation, identifying long-distance paths with high attention weights to support various fine-grained, interpretable herbal medicine recommendation

Table 5 Comparisons between representative AI models in compound toxicity prediction.

TCM-specific applications	Name	Method	Advantages	Authors
Liver Toxicity	-	Multitask Ensemble(NB, LibSVM, KStar, AdaboostM1-J48, Bagging-IBK, J48, and RF)	The first instance of a computational toxicology to screen the hepatotoxic ingredients in TCMs.	He et al. [142]
Toxicity prediction for CM small molecule components	ESTD	DNN, RF, and GBDT	For quantitative toxicity analysis and prediction of small molecular.	Wu et al. [143]
Toxicity and Side effects for COVID-19 treatment	-	OSPF, ANN	An advanced approach to the evaluation of TCM prescriptions for treating flu-like diseases.	Liu et al. [124]
hERG Cardiotoxicity	DeepHIT	DNN/GCN	Multi-model architectures are utilized to enhance the predictive accuracy of cardiotoxicity, hepatotoxicity, and genotoxicity assessments in CM.	Ryu et al. [145]
Liver Toxicity/hERG Cardiotoxicity	FP-ADMET	RF	Generalizable molecular fingerprinting method to predict the toxicity of CM components.	Venkatraman et al. [146]
Liver Toxicity/hERG Cardiotoxicity	ResNet18DNN	DNN	DNNs leverage representation learning on the molecular structures of Chinese medicine (CM) components to accurately predict toxicity endpoints, including carcinogenicity and hepatotoxicity.	Chen et al. [147]

Note, OSPF: ontology-based side-effect prediction framework; NB: Naive Bayes; RF: Random Forest; DNN: Deep Neural Networks; GCN: Graph Convolutional Networks; ANN: Artificial Neural Networks; GBDT: Gradient Boosting Decision Tree

tasks. In practice, the model recommends dried ginger to warm the middle-jiao and dispel cold based on the metapath "symptom of vomiting". Practical applications demonstrate that MGAT excels in explaining mechanisms of TCMFs and provides superior interpretability.

3. Challenges and prospects

The rapid advancement in artificial intelligence (AI) has provided a significant impetus to the innovation and modernization of traditional Chinese medicine, offering new perspectives on studies of TCMFs. However, the AI-assisted research on molecular mechanism of TCMFs still faces many challenges, calling for the development of innovative approaches to address these issues.

3.1. Standardizing TCM data for AI model construction

The incorporation of AI and TCM faces some obstacles about data scarcity and complexity [151]. Effective training of AI models depends on access to large-scale, well-structured, and standardized datasets, whereas datasets in TCM research are often fragmented, heterogeneous, and lacking in uniform data standards. While comprehensive, high-quality datasets typically produce more reliable AI models, smaller, unrepresentative datasets tend to result in overfitting [152]. Current TCM-related datasets suffer from several limitations, including insufficient volume, incomplete types, and poor quality [153]. Besides, existing TCM databases exhibit limited data variety, non-standardized protocols, and significant redundancy. A thorough analysis by Wang et al. identified notable inconsistencies in medicinal herb and ingredient annotations across different databases, particularly regarding TCMF structural information [154]. Additionally, most Chinese medicine databases lack crucial molecular fingerprint data for TCM components and high-throughput transcriptomic data for compounds, herbs, and formulas. To address the scarcity and quality issues of TCM datasets, NER algorithms are introduced for automatic entity identification instead of manual extraction. TCMNER, a NER-based model, has been implemented for the construction of TCM entity dataset [140]. It employs a word–character integrated self-attention mechanism to effectively manage variable-length entities, yielding enhanced character-level representations. Similarly, a BERT–BiLSTM–CRF-based NER framework has been applied to TCM terminology identification [155], combining transfer learning techniques with classical machine learning algorithms. Data augmentation techniques generate synthetic samples to expand and diversify existing datasets. Veit Sandfort's team proposed a CycleGAN-based approach that integrates spatial feature transformation into the generative adversarial network (GAN) framework, thereby enhancing image segmentation performance in scenarios with limited CT imaging data [156]. These approaches expand both the quantity and diversity of training data available to ML algorithms, thereby enhancing predictive accuracy and reliability. In TCM research, studies often face challenges such as limited sample sizes and highly heterogeneous data types. To address these limitations, data augmentation techniques are utilized to generate diverse samples, including molecular graphs of ingredients or targets [157–158], This approach facilitates the development of more robust models with enhanced generalization capabilities. Additionally, this strategy improves classification and recognition performance, enabling models to differentiate more precisely among the components and ingredients of TCMFs.

In the future, data standardization represents the fundamental prerequisite for AI model construction in TCM research. TCM data standardization necessitates a comprehensive approach across multiple dimensions. This includes standardizing terminology for herbal origins, decoction processing, and formula compatibility (e.g., establishing a formal framework for "sovereign-minister-assistant-envoy"compatibility relationships), as well as standardized representations of chemical component structures, target protein sequences, and PK parameters at the molecular level. Furthermore, it requires quantified definitions for syndrome classification and efficacy evaluation indicators in clinical data. Standardized TCM data provides an essential foundation for pharmacological investigation and facilitates the effective implementation of component-based research paradigms for TCMFs.

3.2. Developing interpretable models for the mechanistic investigation of TCMFs

DL is frequently characterized as a black-box model due to its non-interpretable internal structure and computational processes regarding input-output relationships. Despite their strong performance, these models' lack of interpretability prevents domain researchers from understanding their decision-making processes, diminishing trust in their reasoning, particularly in TCM contexts requiring clear elucidation of pharmacological mechanisms and component interactions. Consequently, developing interpretable AI models for TCM is crucial for future advancement. These models have the potential to enhance reliability and explainability while offering deeper insights into the mechanisms of TCM. Some researchers improve AI model interpretability by leveraging knowledge graphs (KGs), which effectively integrate heterogeneous and discrete data from the TCM domain, establishing structured relationships between data entities. [135, 159]

(Fig. 4). For instance, the DL architecture combined with KGs can effectively integrate relationships between TCM and clinical symptoms, enabling more comprehensive node feature representation and providing holistic, higher-order feature information for TCM target prediction [160–161]. Moreover, KG-embedded models, such as KGCN [162] and KGAT [163], offer valuable insights into compound-target interactions through more accurate node representations, achieving enhanced prediction performance. Zhuo et al. developed a KG relational path network (KGRN) for target prediction in TCMFs, combining KG embeddings and GNNs to extract semantic and structural information, capturing formula-target correlations [164]. KGRN emphasizes relational dependencies in paths during information aggregation, enhancing formula-target relationship prediction. The DL-based formula target prediction framework, HTINet2, comprises three key modules: TCM and clinical KG embedding, residual graph representation learning, and supervised target prediction [165]. Its effectiveness derives from two primary factors: the constructed KG contains comprehensive TCM clinical treatment information, providing precise herb and target embedding vectors; and the novel residual graph effectively integrates complex biomolecular characteristics, capturing sophisticated herb-target interactions.

Building upon innovations in KGs and the integration of TCM theoretical frameworks, future interpretable TCM models should address several developmental directions. First, incorporating domain-specific interpretability architectures grounded in TCM foundational theories—such as translating the “sovereign-minister-assistant-envoy” compatibility principle [166] into algorithmic constraints for target prediction and mechanism analysis [167–168] —establishes theory-compatible explanatory frameworks. These frameworks connect data-driven statistical patterns with TCM’s multi-component/multi-target synergy logic. Second, addressing the limitations of text-centric KGs requires the development of multi-modal representations. This approach incorporates herbal imagery (morphological/microscopic features), chemical spatial conformations, and clinical imaging biomarkers (e.g., tongue/pulse diagnostics) [169–170]. Such integration captures TCM’s inherent multidimensionality and facilitates clinically grounded AI applications. Finally, integrating KG symbolic reasoning, DL numerical modeling, and expert validation into a theory-data-expert tripartite framework may enhance the analysis of TCM’s complex mechanisms.

In conclusion, developing AI models with KGs represents an important direction for intelligent TCM research, offering interpretable and comprehensible models for mechanistic investigation of TCMFs.

3.3. Building up graph-based AI models for TCM multi-omics data integration

Multi-omics approaches, including genomics, transcriptomics, proteomics, and metabolomics, utilize high-throughput detection technologies to comprehensively examine molecular profiles across various levels. These methods reveal the complex signatures of physiological processes. The holistic, dynamic, and spatiotemporally specific characteristics of multi-omics align with the comprehensive therapeutic principles of TCM, offering a critical perspective for understanding the molecular mechanisms underlying TCMFs [171]. Through AI integration of multi-omics data, researchers can better understand the multi-component, multi-target, and multi-pathway mechanisms of TCMFs. Current multi-omics data integration methods primarily represent multi-layered information within omics datasets through biological networks. These networks illustrate relationships between cells, molecules, and their interactions. For instance, Sun et al. developed the Scissor algorithm, which determines phenotype-cell associations through a network-regularized sparse regression model on correlation matrices, integrating large-scale phenotype and single-cell expression data [172].

The application of graph-based AI models for TCM multi-omics data integration represents a crucial approach with significant potential for advancing TCM understanding and implementation in modern healthcare. These models effectively capture complex interactions between components, targets, and pathways by representing biological relationships as graphs, providing comprehensive insight into TCMFs’ mechanisms of action.

The integration of multi-omics data, encompassing genomics, proteomics, metabolomics, and transcriptomics, enables researchers to elucidate the “multi-component, multi-target, and multi-pathway” characteristics of TCMFs. Graph-based AI models identify essential nodes (e.g., active compounds or critical pathways) and modules within these networks, facilitating novel therapeutic target discovery and optimizing formula compositions for enhanced efficacy. This approach supports precision medicine through personalized treatment strategies based on individual molecular profiles. Graph-based AI models guide new drug development with well-defined mechanisms and reduced side effects by predicting drug-target interactions and identifying biomarkers associated with TCMFs’ effects.

Furthermore, the integration of multi-omics data through graph-based AI facilitates the connection between traditional medicine and modern science, advancing TCM standardization and internationalization. This approach enhances understanding

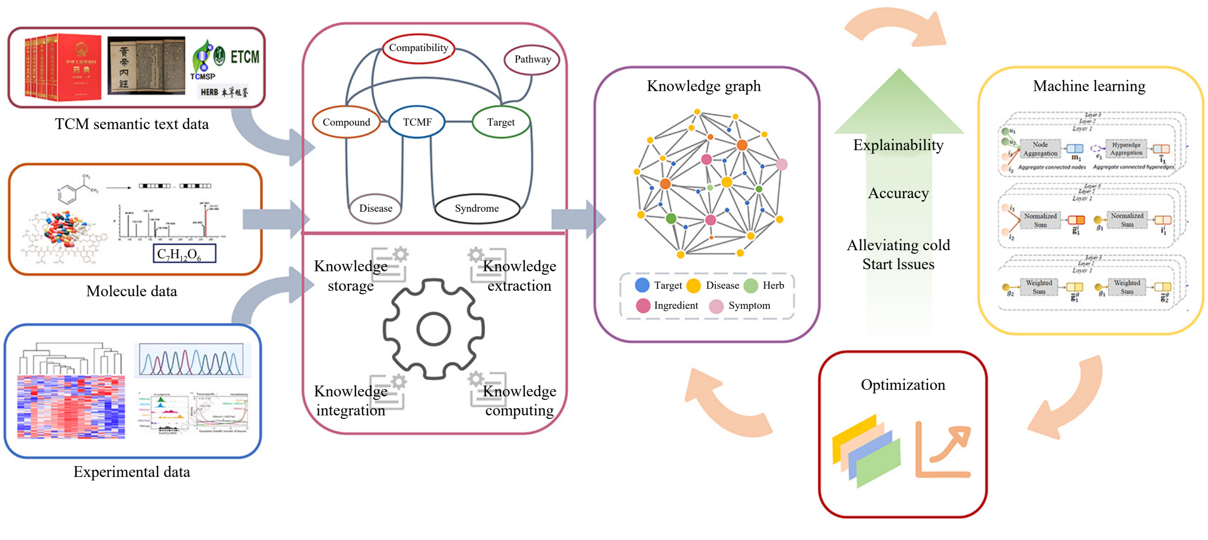

Fig. 4 Knowledge graph plays an important role in AI models of TCM.

of TCMFs' molecular foundations while accelerating the translation of TCM into evidence-based therapies for global health applications.

4. Conclusion

This paper presents a comprehensive review of AI applications in elucidating the mechanisms of TCMFs. It examines practical implementations of AI models in TCM research, including ADME prediction, target identification, synergistic effect characterization, pharmacological mechanism exploration, and formula optimization and recommendation. Collectively, these studies highlight AI's potential to discover novel targets in TCMFs, elucidate synergistic interactions among multi-components, and refine composition of compound formulas. However, current AI-assisted mechanism research of TCMFs remains in its early stages and still faces many challenges, including poor data quality, insufficient algorithm transparency, and limited application scenarios. To deepen the mechanistic investigation of TCMFs through AI models, future efforts should prioritize improving dataset quantity and quality in accordance with TCM data standards, enhancing model interpretability *via* knowledge graphs and developing graph-based method for omics data integration. Applying advanced AI technologies in the research of molecular mechanisms of TCMFs accelerates the modernization and globalization of Chinese medicine.

Funding

This work was supported by the National Key R&D Program of China (No. 2022YFC3502005), the three-year Action Plan for Shanghai TCM Development and Inheritance Program [No. ZY(2021-2023)-0401] and the National Natural Science Foundation of China (No. 82104521).

Declarations Conflict of interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

1 Soon WW, Hariharan M, Snyder MP. High-throughput sequencing for biology and medicine. *Mol Syst Biol*. 2013;9:640. https://doi.org/10.1038/msb.2012.61.

2 Tu YY. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat Med*. 2011;17(10):1217-1220. https://doi.org/10.1038/nm.2471.

3 Wang WY, Xie Y, Zhou H, et al. Contribution of traditional Chinese medicine to the treatment of COVID-19. *Phytomedicine*. 2021;85:153279. https://doi.org/10.1016/j.phymed.2020.153279.

4 Huang K, Zhang P, Zhang ZH, et al. Traditional Chinese medicine (TCM) in the treatment of COVID-19 and other viral infections: efficacies and mechanisms. *Pharmacol Ther*. 2021;225:107843. https://doi.org/10.1016/j.pharmthera.2021.107843.

5 Zhang BL, Wang YYY. The fundamental study on critical scientific problem of prescriptions. *Chin J Nat Med*. 2005;5:258-261. https://lib.cqvip.com/Qikan/Article/Detail?id=20208413.

6 Zhang WD. *Systems biology and modern research on traditional Chinese medicine formulas*. Beijing, BJ: China Science Publishing & Media Ltd.; 2017.

7 Guo DA, Ye M, Wu WY, et al. Systematic analytical methods for complex systems of traditional Chinese medicine and their applications in standards development. *Chin Nat Med Res Dev Forum*. 2013;53-56.

8 Galke L, Ram Y, Raviv L. Deep neural networks and humans both benefit from compositional language structure. *Nat Commun*. 2024;15(1):10816. https://doi.org/10.1038/s41467-024-55158-1.

9 Jamali M, Grannan B, Cai J, et al. Semantic encoding during language comprehension at single-cell resolution. *Nature*. 2024;631(8021):610-616. https://doi.org/10.1038/s41586-024-07643-2.

10 Feuerriegel S, Maarouf A, Bär D, et al. Using natural language processing to analyse text data in behavioural science. *Nat Rev Psychol*. 2025;4:96-111. https://doi.org/10.1038/s44159-024-00392-z.

11 Khurana D, Koli A, Khatter K, et al. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl*. 2022;82:3713-3744. https://doi.org/10.1007/s11042-022-13428-4.

12 Lewis M, Liu Y, Goyal N, et al. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *ACL 2020*. 2020:7871-7880. https://doi.org/10.18653/v1/2020.acl-main.703.

13 Schmidhuber J. Learning to control fast-weight memories: an alternative to dynamic recurrent networks. *Neural Comput*. 1992;4(1):131-139. https://doi.org/10.1162/neco.1992.4.1.131.

14 Pannone A, Raj A, Ravichandran H, et al. Robust chemical analysis with graphene chemosensors and machine learning. *Nature*. 2024;634(8034):572-578. https://doi.org/10.1038/s41586-024-08003-w.

15 Watson JL, Juergens D, Bennett NR, et al. *De novo* design of protein structure and function with RFdiffusion. *Nature*. 2023;620(7976):1089-1100. https://doi.org/10.1038/s41586-023-06415-8.

16 Graves L, Nagisetty V, Ganesh V. Amnesiac machine learning. *Proc AAAI Conf Artif Intell*. 2021;35(13):11516-11524. https://doi.org/10.1609/aaai.v35i13.17371.

17 Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science*. 2015;349:255-260. https://doi.org/10.1126/science.aaa8415.

18 Biamonte J, Wittek P, Pancotti N, et al. Quantum machine learning. *Nature*. 2017;549:195-202. https://doi.org/10.1038/nature23474.

19 Zhao X., Wang L., Zhang YF. et al. A review of convolutional neural networks in computer vision. *Artif Intell Rev*. 2024;57:99. https://doi.org/10.1007/s10462-024-10721-6.

20 M Laad, R Maurya, N Saiyed. Unveiling the vision: a comprehensive review of computer vision in AI and ML.In: 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems. 2024;1-6. https://doi.org/10.1109/ADICS58448.2024.10533631.

21 Voulodimos A, Doulamis N, Doulamis A, et al. Deep learning for computer vision: a brief review. *Comput Intell Neurosci*. 2018;7068349:13. https://doi.org/10.1155/2018/7068349.

22 Salin E, Farah B, Ayache S, et al. Are vision-language transformers learning multimodal tepresentations? A probing Perspective. *Proc Conf AAAI Artif Intell*. 2022;36(10):11248-11257. https://doi.org/10.1609/aaai.v36i10.21375.

23 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60(6):84-90. https://doi.org/10.1145/3065386.

24 Marr D. Visual information processing: the structure and creation of visual representationsPhil. *Phil Trans Soc Lond B*. 1980;290:199-218. https://doi.org/10.1098/rstb.1980.0091.

25 Editorial. Will generative AI transform robotics? *Nat Mach Intell*. 2024;6(6):579-579. https://doi.org/10.1038/s42256-024-00862-2.

26 Kim T, Hong I, Im S, et al. Wing-strain-based flight control of flapping-wing drones through reinforcement learning. *Nat Mach Intell*. 2024;6:992-1005 . https://doi.org/10.1038/s42256-024-00893-9.

27 Zhang KT, Chermprayong P, Xiao F, et al. Aerial additive manufacturing with multiple autonomous robots. *Nature*. 2022;609(7928):709-717. https://doi.org/10.1038/s41586-022-04988-4.

28 de Croon GCHE, De Wagter C, Seidl T. Enhancing optical-flow-based control by learning visual appearance cues for flying robots. *Nat Mach Intell*. 2021;3(1):33-41. https://doi.org/10.1038/s42256-020-00279-7.

29 Chen AI, Balter ML, Maguire TJ, et al. Deep learning robotic guidance for autonomous vascular access. *Nat Mach Intell*. 2020;2(2):104-115. https://doi.org/10.1038/s42256-020-0148-7.

30 Paul D, Sanap G, Shenoy S, et al. Artificial intelligence in drug discovery and development. *Drug Discov Today*. 2021;26(1):80-93. https://doi.org/10.1016/j.drudis.2020.10.010.

31 Gupta RR. Application of artificial intelligence and machine learning in drug discovery. Methods Mol Biol. Heidelberg: Humana Press; 2022: 113-124. https://doi.org/10.1007/978-1-0716-1787-8_4.

32 Sarkar C, Das B, Rawat VS, et al. Artificial intelligence and machine learning technology driven modern drug discovery and development. *Int J Mol Sci*. 2023;24(3):2026. https://doi.org/10.3390/ijms24032026.

33 Rashid MBMA. Artificial intelligence effecting a paradigm shift in drug development. *SLAS Technology*. 2021;26(1):3-15. https://doi.org/10.1177/2472630320956931.

34 Recanatini M, Cabrelle C. Drug research meets network science: where are we? *J Med Chem*. 2020;63(16):8653-8666. https://doi.org/10.1021/acs.jmedchem.9b01989.

35 Zhang P, Wang F, Hu JY, et al. Label propagation prediction of drug-drug interactions based on clinical side effects. *Sci Rep*. 2015;5(1):12339. https://doi.org/10.1038/srep12339.

36 Shalaby KS, Rao SG, Schultz B, et al. SynDRep: a knowledge graph-enhanced tool based on synergistic partner prediction fordrug repurposing. *Bioinform Adv*. 2024; 5(1):vbaf092. https://doi.org/10.1101/2024.08.13.607713.

37 Islam MK, Amaya-Ramirez D, Maigret B, et al. Molecular-evaluated and explainable drug repurposing for COVID-19 using ensemble knowledge graph embedding. *Sci Rep*. 2023;13(1):1-15. https://doi.org/10.1038/s41598-022-26890-9.

38 Mohammed MA, Abdulkareem KH, Dinar AM, et al. Rise of deep learning clinical applications and challenges in omics data: a systematic review. *Diagnostics*. 2023;13(4):664. https://doi.org/10.3390/diagnostics13040664.

39 Ii B. The power of deep learning to ligand-based novel drug discovery. *Expert Opin Drug Discov*. 2020;15(7): 755-764. https://doi.org/10.1080/17460441.2020.1745183.

40 Kuenzi BM, Park J, Fong SH, et al. Predicting drug response and synergy using a deep learning model of human cancer cells. *Cancer Cell*. 2020;38(5):672-684. e676. https://doi.org/10.1016/j.ccell.2020.09.014.

41 Exscientia. Exscientia announces first-in-human study for bristol myers squibb in-licensed PKC theta inhibitor, EXS4318. Accessed 4 November 2024. https://investors.exscientia.ai/press-releases/press-release-details/

2023/Exscientia-Announces-First-in-Human-Study-for-Bristol-Myers-Squibb-In-Licensed-PKC-Theta-Inhibitor-EXS4318/default.aspx.

42 Chakraborty S, Chopra H, Akash S, et al. Artificial intelligence (AI) is paving the way for a critical role in drug discovery, drug design, and studying drug–drug interactions – correspondence. *Int J Surg*. 2023;109(10):3242. https://doi.org/10.1097/JS9.0000000000000564.

43 Kuru HI, Tastan O, Cicek AE. MatchMaker: a deep learning framework for drug synergy prediction. *IEEE/ACM Trans Comput Biol*. 2022;19(4):2334-2344. https://doi.org/10.1109/TCBB.2021.3086702.

44 Zhang TY, Zhang LW, Payne PRO, et al. Synergistic drug combination prediction by integrating multiomics data in deep learning models. Translational Bioinformatics for Therapeutic Development. New York, NY: Springer US. 2021: 223-238. https://doi.org/10.1007/978-1-0716-0849-4_12.

45 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.

46 Ke ZP, Liu MX, Liu J, et al. The application of artificial intelligence in the research and development of traditional Chinese medicine. *Int J Drug Discov Pharm*. 2024; 2024: 100001. https://doi.org/10.53941/ijddp.2024.100001.

47 Song ZL, Chen GX, Chen CYC. AI empowering traditional Chinese medicine? *Chem Sci*. 2024;15(41):16844-16886. https://doi.org/10.1039/D4SC04107K.

48 Liu M, Wang X, Zhou L, et al. Study on extraction and recognition of traditional Chinese medicine tongue manifestation: based on deep learning and migration learning. *J Tradit Chin Med*. 2019;60(10):835-840. https://doi.org/10.13288/j.11-2166/r.2019.10.007.

49 Song C, Wang B, Xu JT. A tongue image classification method based on deep transfer learning. *Comput Sci Eng*. 2021;43(8):1488-1496. https://doi.org/10.3969/j.issn.1007-130X.2021.08.018.

50 Lin Y. Facial complexion classification of traditional Chinese medicine based on statistical features and mobileViT. In: IEEE Inf Technol Netw Electr Autom Control Conf. 2021;7:50-54. https://doi.org/10.1109/itnec60942.2024.10733073.

51 Li MX, Xu R, Zhang T, et al. A retinal vessel segmentation method based on super-pixel and generative adversarial networks. *J Integr Technol*. 2020; 9(6):21-28. https://doi.org/10.12146/j.issn.2095-3135.20200707001.

52 Du FZ, He YJ, Feng XB, et al. Research on TCM syndrome classification Algorithm based on artificial intelligence. *Acta Sci Nat Univ Nankaiensis*. 2023;56(2):12-16. https://doi.org/CNKI:SUN:NKDZ.0.2023-02-001.

53 Zeng ZL. Construction of knowledge map of Yin deficiency syndrome based on ancient books and research on key technologies. Beijing: China Academy of Chinese Medical Sciences. 2023. https://doi.org/10.27658/d.cnki.gzzyy.2022.000035.

54 Kuang H. Construction of knowledge graph in treatise on febrile diseases and its application in Yang Ming disease. Zhejiang: Zhejiang Chinese Medical University. 2022. https://doi.org/10.27465/d.cnki.gzzyc.2021.000001.

55 Bender A, Glen RC. Molecular similarity: a key technique in molecular informatics. *Org Biomol Chem*. 2004;2(22):3204-3218. https://doi.org/10.1039/b409813g.

56 Mackenzie PI, Somogyi AA, Miners JO. Advances in drug metabolism and pharmacogenetics research in Australia. *Pharmacol Res*. 2017;116:7-19. https://doi.org/10.1016/j.phrs.2016.12.008.

57 Kirchmair J, Göller AH, Lang D, et al. Predicting drug metabolism: experiment and/or computation? *Nat Rev Drug Discovery*. 2015;14(6):387-404. https://doi.org/10.1038/nrd4581.

58 Sun L, Mi K, Hou Y, et al. Pharmacokinetic and pharmacodynamic drug–drug Interactions: research methods and applications. *Metabolites*. 2023;13(8):897. https://doi.org/10.3390/metabo13080897.

59 Zhang X, Liu T, Fan X, et al. *In silico* modeling on ADME properties of natural products: classification models for blood-brain barrier permeability, its application to traditional Chinese medicine and *in vitro* experimental validation. *J Mol Graphics Modell*. 2017;75:347-354. https://doi.org/10.1016/j.jmgm.2017.05.021.

60 Guo FF, Tang X, Zhang W, et al. Exploration of the mechanism of traditional Chinese medicine by AI approach using unsupervised machine learning for cellular functional similarity of compounds in heterogeneous networks, XiaoErFuPi Granules as an example. *Pharmacol Res*. 2020;160:105077. https://doi.org/10.1016/j.phrs.2020.105077.

61 Wang P, Li K, Tao Y, et al. TCM-ADMEpred: a novel strategy for poly-pharmacokinetics prediction of traditional Chinese medicine based on single constituent pharmacokinetics, structural similarity, and mathematical modeling. *J Ethnopharmacol*. 2019;236:277-287. https://doi.org/10.1016/j.jep.2018.07.008.

62 Zhang W, Huai Y, Miao Z, 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/10.3389/fphar.2019.00743.

63 Feinberg EN, Joshi E, Pande VS, et al. Improvement in ADMET prediction with multitask deep featurization. *J Med Chem*. 2020;63(16):8835-8848. https://doi.org/10.1021/acs.jmedchem.9b02187.

64 Duvenaud D, Maclaurin D, Aguilera-Iparraguirre J, et al. Convolutional networks on graphs for learning molecular fingerprints. *Adv Neural Inf Process Syst* . 2015;2:2224-2232.

65 Yang K, Swanson K, Jin W, et al. Analyzing learned molecular representations for property prediction. *J Chem Inf Model*. 2019;59(8):3370-3388. https://doi.org/10.1021/acs.jcim.9b00237.

66 Liu K, Sun X, Jia L, et al. Chemi-Net: a molecular graph convolutional network for accurate drug property prediction. *Int J Mol Sci*. 2019;20(14): 3389. https://doi.org/10.3390/ijms20143389.

67 Hou TJ, Wang JM, Zhang W, et al. ADME evaluation in drug discovery. 7. Prediction of oral absorption by correlation and classification. *J Chem Inf Model*. 2007;47(1):208-18. https://doi.org/10.1021/ci600343x.

68 Rodríguez-Pérez R, Trunzer M, Schneider N, et al. Multispecies machine learning predictions of *in vitro* intrinsic clearance with uncertainty quantification analyses. *Mol Pharm*. 2023;20:383. https://doi.org/10.1021/acs.molpharmaceut.2c00680.

69 Menon D, Sharma AM, Ani R. Computational approaches for drug discovery: a review of artificial intelligence techniques for ADME/T prediction and virtual screening. *Int Conf Comput Commun Netw Technol*. 2024; 2024: 1-11. https://doi.org/10.1109/icccnt61001.2024.10724127.

70 Wishart DS, Knox C, Guo AC, et al. DrugBank: a comprehensive resource for *in silico* drug discovery and exploration. *Nucleic Acids Res*. 2006; 34 (Database issue): D668-D672. https://doi.org/10.1093/nar/gkj067.

71 Ru J, Li P, Wang J,et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminf*. 2014; 6(1): 13. https://doi.org/10.1186/1758-2946-6-13.

72 Pires DEV, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem*. 2015;58(9):4066-4072. https://doi.org/10.1021/acs.jmedchem.5b00104.

73 Tao W, Li B, Gao S, et al. CancerHSP: anticancer herbs database of systems pharmacology. *Sci Rep*. 2015;5:11481. https://doi.org/10.1038/srep11481.

74 Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep*. 2017;7:42717. https://doi.org/10.1038/srep42717.

75 Xiong GL, Wu ZX, Yi JC, et al. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res*. 2021;49(W1):W5-W14. https://doi.org/10.1093/nar/gkab255.

76 Wang P, Wang S, Chen H, et al. TCMIP v2.0 powers the identification of chemical constituents available in xinglou chengqi Decoction and the exploration of pharmacological mechanisms acting on stroke complicated with tanre fushi syndrome. *Front Pharmacol*. 2021;12:598200. https://doi.org/10.3389/fphar.2021.598200.

77 Fu L, Shi S, Yi J, et al. ADMETlab 3.0: an updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support. *Nucleic Acids Res*. 2024;52(W1):W422-W431. https://doi.org/10.1093/nar/gkae236.

78 Li YL, Lin ZJ, Wang Y, et al. Unraveling the mystery of efficacy in Chinese medicine formula: new approaches and technologies for research on pharmacodynamic substances. *Arabian J Chem*. 2022;15(11):104302. https://doi.org/10.1016/j.arabjc.2022.104302.

79 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.

80 Zhang JH, Zhu Y, Fan XH, et al. Efficacy-oriented compatibility for component-based Chinese medicine. *Acta Pharmacol Sin*. 2015;36(6):654-658. https://doi.org/10.1038/aps.2015.8.

81 Wu Y, Liu XQ, Li GW. Integrated bioinformatics and network pharmacology to identify the therapeutic target and molecular mechanisms of Huangqin Decoction on ulcerative colitis. *Sci Rep*. 2022;12(1): 159. https://doi.org/10.1038/s41598-021-03980-8.

82 Liu Y, Luo D, Xu B. The combination of molecular docking and network pharmacology reveals the molecular mechanism of Danggui Niantong Decoction in treating gout. *Medicine*. 2022;101(47):e31535. https://doi.org/10.1097/MD.0000000000031535.

83 Xue RC, Fang Z, Zhang MX, et al. TCMID: traditional Chinese medicine integrative database for herb molecular mechanism analysis. *Nucleic Acids Res*. 2012;41(D1):D1089-D1095. https://doi.org/10.1093/nar/gks1100.

84 Liu ZY, Guo FF, Wang Y, et al. BATMAN-TCM: a bioinformatics analysis tool for molecular mechanism of traditional Chinese medicine. *Sci Rep*. 2016; 6(1):21146. https://doi.org/10.1038/srep21146.

85 Xu HY, Zhang YQ, Liu ZM, et al. ETCM: an encyclopaedia of traditional Chinese medicine. *Nucleic Acids Res*. 2019;47(D1):D976-D982. https://doi.org/10.1093/nar/gky987.

86 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.

87 Fang SS, Dong L, Liu L, et al. HERB: a high-throughput experiment- and reference-guided database of traditional Chinese medicine. *Nucleic Acids Res*. 2020;49(D1):D1197-D1206. https://doi.org/10.1093/nar/gkaa1063.

88 Tian SS, Zhang JB, Yuan SL, et al. Exploring pharmacological active ingredients of traditional Chinese medicine by pharmacotranscriptomic map in ITCM. *Briefings Bioinf*. 2023;24(2):bbad027. https://doi.org/10.1093/bib/bbad027.

89 Liu XY, Liu JY, Fu BZ, et al. DCABM-TCM: a database of constituents absorbed into the blood and metabolites of traditional Chinese medicine. *J Chem Inf Model*. 2023;63(15):4948-4959. https://doi.org/10.1021/acs.jcim.3c00365.

90 Lv QJ, Chen GX, He HH, et al. TCMBank-the largest TCM database provides deep learning-based Chinese-Western medicine exclusion prediction. *Signal Transduction Targeted Ther*. 2023;8(1):127. https://doi.org/10.1038/s41392-023-01339-1.

91 Yang DQ, Zhu Z, Yao Q, et al. ccTCM: a quantitative component and compound platform for promoting the research of traditional Chinese medicine. *Comput Struct Biotechnol*. 2023; 21: 5807-5817. https://doi.org/10.1016/j.csbj.2023.11.030.

92 Ye YN, Liang DF, Yi JH, et al. IGTCM: an integrative genome database of traditional Chinese medicine plants. *The Plant Genome*. 2023;16(2):e20317. https://doi.org/10.1002/tpg2.20317.

93 Ren ZX, Ren YM, Li ZT, et al. TCMM: a unified database for traditional Chinese medicine modernization and therapeutic innovations. *Comput*

Struct Biotechnol J. 2024;23:1619-1630. https://doi.org/10.1016/j.csbj.2024.04.016.

94 Alexander S, Friedl P. Cancer invasion and resistance: interconnected processes of disease progression and therapy failure. *Trends Mol Med*. 2012;18(1):13-26. https://doi.org/10.1016/j.molmed.2011.11.003.

95 Luo M, Jiao JF, Wang RQ. Screening drug target combinations in disease-related molecular networks. *BMC Bioinf*. 2019;20(Suppl 7):198. https://doi.org/10.1186/s12859-019-2730-8.

96 Mahase E. Breast cancer: NICE approves new drug combination treatment. *BMJ*. 2019; 2019: l4727. https://doi.org/10.1136/bmj.l4727.

97 Peraman R, Sure SK, Dusthackeer VNA, et al. Insights on recent approaches in drug discovery strategies and untapped drug targets against drug resistance. *Future J Pharm Sci*. 2021;7(1):56. https://doi.org/10.1186/s43094-021-00196-5.

98 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.

99 Malik MA, Wani MY, Hashmi AA. Combination therapy: current status and future perspectives. In: Combination Therapy Against Multidrug Resistance Bacteria. Springer; 2020:1-38. https://doi.org/10.1016/B978-0-12-820576-1.00001-1.

100 Pan YC, Ren HT, Lan L, et al. Review of predicting synergistic drug combinations. *Life (Basel)*. 2023;13(9):1878. https://doi.org/10.3390/life13091878.

101 Naveja JJ, Stumpfe D, Medina-Franco JL, et al. Exploration of target synergy in cancer treatment by cell-based screening assay and network propagation analysis. *J Chem Inf Model*. 2019; 59(6):3072-3079. https://doi.org/10.1021/acs.jcim.9b00036.

102 Cheng FX, Kovács IA, Barabási AL. Network-based prediction of drug combinations. *Nat Commun*. 2019;10(1):1-11. https://doi.org/10.1038/s41467-018-07882-8.

103 Zheng SC, Baak JP, Li S, et al. Network pharmacology analysis of the therapeutic mechanisms of the traditional Chinese herbal formula Lian Hua Qing Wen in corona virus disease 2019 (COVID-19), gives fundamental support to the clinical use of LHQW. *Phytomedicine*. 2020;79:153336. https://doi.org/10.1016/j.phymed.2020.153336.

104 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.

105 Wang YY, Yang HB, Chen LX, et al. Network-based modeling of herb combinations in traditional Chinese medicine. *Briefings Bioinf*. 2021;22(5): bbab106. https://doi.org/10.1093/bib/bbab106.

106 Gan X, Shu Z, Wang X, et al. Network medicine framework reveals generic herb-symptom effectiveness of traditional Chinese medicine. *Sci Adv*. 2023;9:eadh0215. https://doi.org/10.1126/sciadv.adh0215.

107 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.

108 Jeon M, Kim S, Park S, et al. *In silico* drug combination discovery for personalized cancer therapy. *BMC Syst Biol*. 2018;12(2):16. https://doi.org/10.1186/s12918-018-0546-1.

109 Ung CY, Li H, Cao ZW, et al. Are herb-pairs of traditional Chinese medicine distinguishable from others? Pattern analysis and artificial intelligence classification study of traditionally defined herbal properties. *J Ethnopharmacol*. 2007;111(2):371-377. https://doi.org/10.1016/j.jep.2006.11.037.

110 Xue Q, Gao B, Wen J, et al. A deep learning-based herb pair discovering approach. *Big Data*. 2022;1496: 135-149. https://doi.org/10.1007/978-981-16-9709-8_10.

111 Yang F, Zhang SJ, Pan W, et al. Signaling repurposable drug combinations against COVID-19 by developing the heterogeneous deep herb-graph method. *Briefings Bioinf*. 2022;23(5):bbac124. https://doi.org/10.1093/bib/bbac124.

112 Preuer K, Lewis RPI, 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.

113 Chen J, Wang YK, Gao Y, et al. Protection against COVID-19 injury by Qingfei Paidu Decoction *via* anti-viral, anti-inflammatory activity and metabolic programming. *Biomed Pharmacother*. 2020;129:110281. https://doi.org/10.1016/j.biopha.2020.110281.

114 Li S, Zhang B, Zhang NB. Network target for screening synergistic drug combinations with application to traditional Chinese medicine. *BMC Syst Biol*. 2011;5(1):S10. https://doi.org/10.1186/1752-0509-5-10.

115 Huang L, Li FH, Sheng JT, et al. DrugComboRanker: drug combination discovery based on target network analysis. *Bioinformatics*. 2014;30(12): i228-i236. https://doi.org/10.1093/bioinformatics/btu278.

116 Sun BZ, Liu CX, Li YL, et al. An approach to co-medication mechanism mining of Chinese Materia Medica and western medicines based on complex networks with the multi-source heterogeneous information. *Inf Fusion*. 2024;102:102081. https://doi.org/10.1016/j.inffus.2023.102081.

117 Li X, Liu Z, 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.

118 Zhu X, Yao Q, Yang P, 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.

119 Xiong H, Li N, Zhao LQQ, et al. Integrated serum pharmacochemistry, metabolomics, and network pharmacology to reveal the material basis and mechanism of Danggui Shaoyao San in the treatment of primary dysmenorrhea. *Front Pharmacol*. 2022;13:942955. https://doi.org/10.3389/fphar.2022.942955.

120 Lu X, Ma W, Fan B, et al. Integrating network pharmacology, transcriptome and artificial intelligence for investigating into the effect and mechanism of Ning Fei Ping Xue Decoction against the acute respiratory distress syndrome. *Front Pharmacol*. 2021;12:731377. https://doi.org/10.3389/fphar.2021.731377.

121 Jiang YP, Zheng YF, Dong Q, et al. Metabolomics combined with network pharmacology to study the mechanism of Shentong Zhuyu Decoction in the treatment of rheumatoid arthritis. *J Ethnopharmacol*. 2022;285:114846. https://doi.org/10.1016/j.jep.2021.114846.

122 Ma C, Zhao J, Zheng GY, et al. Qijiao Shengbai Capsule alleviated leukopenia by interfering leukotriene pathway: integrated network study of multi-omics. *Phytomedicine*. 2024;128:155424. https://doi.org/10.1016/j.phymed.2024.155424.

123 Ruchawapol C, Fu WW, Xu HX. A review on computational approaches that support the researches on traditional Chinese medicines (TCM) against COVID-19. *Phytomedicine*. 2022;104:154324. https://doi.org/10.1016/j.phymed.2022.154324.

124 Wang ZH, Li L, Song M, et al. Evaluating the traditional Chinese medicine (TCM) officially recommended in china for COVID-19 using ontology-based side-effect prediction framework (OSPF) and deep learning. *J Ethnopharmacol*. 2021;272:113957. https://doi.org/10.1016/j.jep.2021.113957.

125 Zhao J, Tian S, Lu D, et al. Systems pharmacological study illustrates the immune regulation, anti-infection, anti-inflammation, and multi-organ protection mechanism of Qing-Fei-Pai-Du Decoction in the treatment of COVID-19. *Phytomedicine*. 2021;85:153315. https://doi.org/10.1016/j.phymed.2020.153315.

126 Qi GP, Zhang AJ, Xu Z, et al. Application of a complex network modeling approach to explore the material basis and mechanisms of traditional Chinese medicine: a case study of Xuefu Zhuyu Decoction for the treatment of two types of angina pectoris. *IEEE Access*. 2022;10:114103-114117. https://doi.org/10.1109/ACCESS.2022.3217926.

127 Fu Y, Fang YY, Gong S, et al. Deep learning-based network pharmacology for exploring the mechanism of licorice for the treatment of COVID-19. *Sci Rep*. 2023;13(1):5844. https://doi.org/10.1038/s41598-023-31380-7.

128 Li J, Liu H, Yang ZZ, et al. Synergistic effects of cryptotanshinone and senkyunolide I in Guanxinning Tablet against endogenous thrombus formation in zebrafish. *Front Pharmacol*. 2020;11:622787. https://doi.org/10.3389/fphar.2020.622787.

129 Eltager M, Abdelaal T, Mahfouz A, et al. scMoC: single-cell multi-omics clustering. *Bioinf Adv*. 2022; 2(1): vbac011. https://doi.org/10.1093/bioadv/vbac011.

130 Flores JE, Claborne DM, Weller ZD, et al. Missing data in multi-omics integration: recent advances through artificial intelligence. *Front Artif Intell*. 2023; 6:1098308. https://doi.org/10.3389/frai.2023.1098308.

131 Cao ZJ, Gao G. Multi-omics single-cell data integration and regulatory inference with graph-linked embedding. *Nat Biotechnol*. 2022;40(10):1458-1466. https://doi.org/10.1038/s41587-022-01284-4.

132 Pan HW, He SN, Zhang KJ, et al. AMAM: an attention-based multimodal alignment model for medical visual question answering. *Knowledge-Based Syst*. 2022;255:109763. https://doi.org/10.1016/j.knosys.2022.109763.

133 Shen T, Zhang F, Cheng J. A comprehensive overview of knowledge graph completion. *Knowledge-Based Syst*. 2022;255:109597. https://doi.org/10.1016/j.knosys.2022.109597.

134 Yu ZX, Li JT, Chen L, et al. Unifying multi-associations through hypergraph for bundle recommendation. *Knowledge-Based Syst*. 2022;255:109755. https://doi.org/10.1016/j.knosys.2022.109755.

135 Alshahrani M, Almansour A, Alkhaldi A, et al. Combining biomedical knowledge graphs and text to improve predictions for drug-target interactions and drug-indications. *PeerJ*. 2022;10:e13061. https://doi.org/10.7717/peerj.13061.

136 Bonner S, Barrett IP, Ye C, et al. Understanding the performance of knowledge graph embeddings in drug discovery. *Artif Intell Life Sci*. 2022;2: 100036. https://doi.org/10.1016/j.ailsci.2022.100036.

137 Ali M, Hoyt CT, Domingo-Fernández D, et al. BioKEEN: a library for learning and evaluating biological knowledge graph embeddings. *Bioinformatics*. 2019;35(18):3538-3540. https://doi.org/10.1093/bioinformatics/btz117.

138 Zhou C, Cai CP, Huang XT, et al. TarKG: a comprehensive biomedical knowledge graph for target discovery. *Bioinformatics*. 2024;40(10):btae598. https://doi.org/10.1093/bioinformatics/btae598.

139 Zhu X, Gu YM, Xiao ZF. HerbKG: constructing a herbal-molecular medicine knowledge graph using a two-stage framework based on deep dtransfer learning. *Front Genet*. 2022;13:799349. https://doi.org/10.3389/fgene.2022.799349.

140 Liu Z, Luo CY, Zheng ZY, et al. TCMNER and PubMed: a novel Chinese character-level-based model and a dataset for TCM named entity recognition. *J Healthcare Eng*. 2021;2021:3544281. https://doi.org/10.1155/2021/3544281.

141 Cyranoski D. Why Chinese medicine is heading for clinics around the world. *Nature*. 2018;561(7724):448-450. https://doi.org/10.1038/d41586-018-06782-7.

142 He SB, Zhang XL, Lu S, et al. A computational toxicology approach to screen the hepatotoxic ingredients in traditional Chinese medicines: polygonum multiflorum thunb as a case study. *Biomolecules*. 2019;9(10):577. https://doi.org/10.3390/biom9100577.

143 Wu KD, Wei GW. Quantitative toxicity prediction using topology based multitask deep neural networks. *J Chem Inf Model*. 2018;58(2):520-531. https://doi.org/10.1021/acs.jcim.7b00558.

144 Tran TTV, Surya Wibowo A, Tayara H, et al. Artificial intelligence in drug toxicity prediction: recent advances, challenges, and future perspectives. *J Chem Inf Model*. 2023;63(9):2628-2643. https://doi.org/10.1021/acs.jcim.3c00200.

145 Ryu JY, Lee MY, Lee JH, et al. DeepHIT: a deep learning framework for prediction of hERG-induced cardiotoxicity. *Bioinformatics*. 2020;36(10): 3049-3055. https://doi.org/10.1093/bioinformatics/btaa075.

146 Venkatraman V. FP-ADMET: a compendium of fingerprint-based ADMET prediction models. *J Cheminf*. 2021;13(1):75. https://doi.org/10.1186/s13321-021-00557-5.

147 Chen Z, Jiang Y, Zhang XY, et al. ResNet18DNN: prediction approach of drug-induced liver injury by deep neural network with ResNet18. *Briefings Bioinf*. 2021;23(1):bbab503. https://doi.org/10.1093/bib/bbab503.

148 Zhou WA, Yang K, Zeng JY, et al. FordNet: recommending traditional Chinese medicine formula *via* deep neural network integrating phenotype and molecule. *Pharmacol Res*. 2021;173:105752. https://doi.org/10.1016/j.phrs.2021.105752.

149 Zhao W, Lu WK, Li ZY, et al. TCM herbal prescription recommendation model based on multi-graph convolutional network. *J Ethnopharmacol*. 2022;297:115109. https://doi.org/10.1016/j.jep.2022.115109.

150 Jin YY, Ji WD, Shi Y, et al. Meta-path guided graph attention network for explainable herb recommendation. *Health Inf Sci*. 2023;11(1):5. https://doi.org/10.1007/s13755-022-00207-6.

151 Basu T, Wolf S, Menzer O. The ethics of machine learning in medical sciences: where do we stand today? *Indian J Dermatol*. 2020;65:358. https://doi.org/10.4103/ijd.IJD_419_20.

152 Shimao H, Khern-am-nuai W, Kannan K, et al. Strategic best response fairness in fair machine learning. *AIES '22*. 2022;664. https://doi.org/10.1145/3514094.3534194.

153 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.

154 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.

155 Qu QQ, Kan HX, Wu YT, et al. Named entity recognition of TCM text based on bert model. *Proc-Int Forum Electr Eng Autom*. 2020; 2020: 652-655. https://doi.org/10.1109/IFEEA51475.2020.00139.

156 Sandfort V, Yan K, Pickhardt PJ, et al. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Sci Rep*. 2019;9(1):1-9. https://doi.org/10.1038/s41598-018-37186-2.

157 Zou JP, Yu JL, Hu PW, et al. STAGAN: an approach for improve the stability of molecular graph generation based on generative adversarial networks. *Comput Biol Med*. 2023;167:107691. https://doi.org/10.1016/j.compbiomed.2023.107691.

158 Chen YY, Wang ZX, WangL, et al. Deep generative model for drug design from protein target sequence. *J Cheminf*. 2023;15(1):38. https://doi.org/10.1186/s13321-023-00702-2.

159 Taneja SB, Callahan TJ, Paine MF, et al. Developing a knowledge graph for pharmacokinetic natural product-drug interactions. *J Biomed Inf*. 2023;140: 104341. https://doi.org/10.1016/j.jbi.2023.104341.

160 Tsubaki M, Tomii K, Sese J. Compound–protein interaction prediction with end-to-end learning of neural networks for graphs and sequences. *Bioinformatics*. 2019;35(2):309-318. https://doi.org/10.1093/bioinformatics/bty535.

161 Georgousis S, Kenning MP, Xie X. Graph deep learning: state of the art and challenges. *IEEE Access*. 2021;9:22106-22140. https://doi.org/10.1109/ACCESS.2021.3055280.

162 Kojima R, Ishida S, Ohta M, et al. kGCN: a graph-based deep learning framework for chemical structures. *J Cheminf*. 2020;12(1):32. https://doi.org/10.1186/s13321-020-00435-6.

163 Wu ZH, Zhang XL, Lin XL. KGAT: predicting drug-target interaction based on knowledge graph attention network. *ICIC 2022*. 2022;13394: 438-450. https://doi.org/10.1007/978-3-031-13829-4_38.

164 Gong Z, Zhang NX, He J. KGRN: knowledge graph relational path network for target prediction of TCM prescriptions. *ICIC 2021*. 2021; 2021: 148-161. https://doi.org/10.1007/978-3-030-84532-2_14.

165 Duan P, Yang K, Su X, et al. HTINet2: herb–target prediction *via* knowledge graph embedding and residual-like graph neural network. *Briefings Bioinf*. 2024;25(5):bbae414. https://doi.org/10.1093/bib/bbae414.

166 Guo SY, Deng ZH, Li Y, et al. A review of theoretical research on interpretation of scientific connotation of compatibility of traditional Chinese medicine compounds. *Chin J Exp Tradit Med Formulae*. 2024;30(23): 338-344. https://doi.org/10.13422/j.cnki.syfjx.20241645.

167 Amador-Domínguez E, Serrano E, Manrique D. GEnI: a framework for the generation of explanations and insights of knowledge graph embedding predictions. *Neurocomputing*. 2023;521:199-212. https://doi.org/10.1016/j.neucom.2022.12.010.

168 Gaur M, Faldu K, Sheth A. Semantics of the black-box: can knowledge graphs help make deep learning systems more interpretable and explainable? *IEEE Internet Comput*. 2021;25(1):51-59. https://doi.org/10.1109/MIC.2020.3031769.

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

170 Qian YY, Wang XT, Cai LL, et al. Model informed precision medicine of Chinese herbal medicines formulas–A multi-scale mechanistic intelligent model. *J Pharm Anal*. 2024;14(4):100914. https://doi.org/10.1016/j.jpha.2023.12.004.

171 Cheng H, Liu W, Qu Y, et al. Unleashing the potential of decentralized federated learning in healthcare: a comprehensive survey on insights and challenges. *IEEE Int Conf Data Sci Cyberspace*. 2024; 2024: 53-60. https://doi.org/10.1109/DSC63484.2024.00015.

172 Acar A, Aksu H, Uluagac AS, et al. A survey on homomorphic encryption schemes: theory and implementation. *ACM Comput Surv*. 2018; 51(4): 79. https://doi.org/10.1145/3214303.