Proceedings of ICEGEE 2026 Symposium: AI-Based Medicine and Biological Data Analysis
DOI: 10.54254/2753-8818/2026.AU35394

Research Progress in AI-Assisted Drug Discovery and Development from TCM

Tianyu Chen

School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing, China
2020240147@stu.cpu.edu.cn

Abstract. The unique advantage of traditional Chinese medicine (TCM) lies in its overall regulatory properties, which makes it particularly promising in the treatment of complex diseases. However, for a long time, the discovery of TCM drugs has been limited by inefficient, experience-driven methods and experimental screening methods that are difficult to identify active compound groups, which seriously limits the development of targeted therapy based on TCM. The rapid advancement of artificial intelligence today provides a feasible way forward. The rapid advancement of artificial intelligence today provides a feasible way forward. In this paper, the application of AI in the research and development of new drugs of TCM was systematically reviewed, focusing on the three key stages of AI in the research and development of new drugs of TCM : target recognition, active compound screening and efficacy evaluation. By integrating multi-omics data with TCM syndrome databases, artificial intelligence-driven methods can transform empirical prescriptions into quantifiable target-regulatory networks. Virtual screening technology can quickly predict potential active molecules from a large compound library, and ADMET prediction model can evaluate the absorption, distribution, metabolism, excretion and toxicity of candidate drugs. Looking forward to the future, the evolution of the integration of TCM and artificial intelligence is expected to move from single-target to multi-target mechanisms, from experience to algorithm-driven prescription design, from disease treatment to broader state intervention, and ultimately to outline an innovative path for the modernization of TCM.

Keywords: artificial intelligence, traditional Chinese medicine R&D, target identification, virtual screening, ADMET prediction

1. Introduction

TCM adopts the method of holistic view and treats the body as a whole, which has therapeutic value for complex diseases. In contrast, modern biomedicine focuses on targeted therapy, and TCM remains an important source of drug discovery in China [1].

However, due to its complexity, the research of TCM is challenging. Medicinal plants contain a large number of compounds that act on multiple body systems, making the identification of key active ingredients very difficult. TCM also emphasizes the synergy between components, which is different from the reductionist method of modern science. In addition, the treatment method can change the chemical properties and further complicate the analysis.

DOI: 10.54254/2753-8818/2026.AU35394

Artificial intelligence (AI) has recently accelerated drug discovery. For example, Inke Intelligence developed a method to treat idiopathic pulmonary fibrosis within 18 months [2], which is much faster than traditional methods, demonstrating the efficiency and potential of AI [3].

This paper studies the application of artificial intelligence in the design of TCM drugs, focusing on target recognition, active ingredient screening and effect evaluation. The aim is to show how AI integrates TCM and modern medicine to promote the development of innovative drugs [4].

2. Literature review

2.1. Overview of traditional Chinese medicine

TCM has a long and rich history in China. In clinical practice, Chinese medicine practitioners mainly rely on a method known as syndrome differentiation to guide treatment ; this means that it is necessary to tailor the treatment plan according to the overall situation of the patient. The medicinal materials used in TCM are mainly derived from plants, animals or minerals. Among them, medicinal materials derived from plants are the most common, which is also the reason why TCM is sometimes called herbal medicine. The active ingredients in TCM are very complex. A single herb can contain hundreds or even thousands of chemical components, covering a wide range of structural categories. The biological targets involved in these compounds are also diverse. In diseases such as cancer, the targets of TCM components go beyond traditional enzymes and receptors, including ion channels, transporters, transcription factors, nucleic acids (DNA/RNA), epigenetic regulators and microbiome [5]. This multi-component and multi-target feature poses a serious obstacle to modern drug development.

To Overcome this bottleneck requires a new method framework that can handle large-scale data, model complex networks, and provide reliable predictions [6]. This is precisely what artificial intelligence has become not only useful, but also necessary.

2.2. Overview of artificial intelligence and drug development

Artificial Intelligence (AI) is a computer algorithm designed to perform tasks that usually require human intelligence. By learning and recognizing patterns from data, artificial intelligence systems expand the capabilities of machines in areas such as data processing and pattern recognition [7]. The core technologies of artificial intelligence are machine learning (ML) and deep learning (DL), including artificial neural network (ANN), Bayesian network, convolutional neural network (CNN) [8], recurrent neural network (RNN) and other models.

With the rapid advancement of computing power and the maturity of big data technology, AI has been gradually applied to various stages of computer-aided drug discovery [9], including drug design, formulation prediction and formulation research [10]. In the field of drug design, AI has achieved remarkable results in predicting drug activity, analyzing molecular conformation and designing synthetic routes by combining quantitative structure-property / activity relationship (QSAR) model and molecular docking technology [11].

Taking into account the above advantages and challenges of the modernization of TCM, as well as the unique potential of artificial intelligence, the concept of applying artificial intelligence technology to the key stage of TCM drug design came into being. Specifically, this application can be broken down into three key stages: identification and characterization of targets, screening of active ingredients, and efficacy evaluation tests. The following sections will elaborate on the application of artificial intelligence in these three stages.

DOI: 10.54254/2753-8818/2026.AU35394

3. Applications of artificial intelligence in TCM research and development

This section provides a systematic overview of the specific applications of artificial intelligence in the three key stages of TCM R&D, including target identification and characterisation, screening of bioactive compounds, and pharmacological efficacy testing.

3.1. Target identification and characterisation

Target identification and characterisation is the primary stage in the R&D of new TCM drugs and one of the most critical steps. Traditional target discovery relies on extensive experimental screening and biological validation, which is characterised by long cycles, high costs and low success rates. The introduction of artificial intelligence technology has provided a novel solution for target identification; three specific prediction approaches are outlined below.

Firstly, Multi-omics-integrated target prediction. By integrating multi-omics data (genomics, proteomics, metabolomics) with databases of TCM syndromes, artificial intelligence transforms empirical compatibility into quantifiable target regulatory networks [12].The study utilised multi-layer perceptrons, decision tree regression and gradient boosting regression algorithms to construct a multi-target, multi-pharmacology prediction model (mTPP), successfully identifying 20 candidate drugs with potential effects on drug-induced liver injury [13] .

Second, knowledge graph construction based on literature mining. The researchers established a knowledge map of TCM-Syndrome-Target by mining classical texts and modern research literature. A two-stage transfer learning model was introduced to generate TCM prescriptions from unstructured resources and medical knowledge, linking TCM theory with modern biomedical data [14].

Thirdly, the prediction of herb-component-target interaction based on deep learning. Graph neural networks (GNNs) can be used to analyze the interaction between TCM compound components and disease-related proteins, thereby predicting potential therapeutic targets. For example, GNN analysis predicted that baicalin exerts anti-inflammatory effects by inhibiting the IL-6 / STAT3 pathway.

These artificial intelligence-based methods accelerate the target discovery process and provide the necessary prerequisites for narrowing the scope of candidate drug screening.

3.2. Screening of active ingredients

The screening of bioactive compounds is a core component of the research and development of new drugs of TCM. In the face of the vast chemical space of TCM, artificial intelligence-assisted virtual screening technology helps to identify its complex components and conduct rapid and accurate screening [8].

Virtual screening is a computational method that uses computer simulation technology to quickly and economically predict and screen potential candidate molecules from a massive compound library. These molecules may have the required biological activity for specific biological targets. A literature systematically reviewed the application status and future prospects of machine learning in natural products and molecules of TCM, and clarified that target screening based on pathological mechanism is the key direction of natural product research [15].

A machine learning-based drug reuse framework was proposed to identify new therapeutic applications of herbal compounds, demonstrating the potential of computational methods in expanding the clinical application of natural products [16]. The challenge of large-scale herb target recognition was addressed by reverse docking [17]. This method uses Autodock Vina for structure-

DOI: 10.54254/2753-8818/2026.AU35394

based virtual screening, and has achieved significant improvement in target prediction accuracy. During the COVID-19 pandemic, a team used machine learning and pharmacophore modeling to perform virtual screening of Indonesian herbal compounds to evaluate their potential as adjuvant therapy for COVID-19 [18]. In different applications, machine learning was used to evaluate herbal extracts as potential vaccine adjuvants, thereby simplifying the candidate selection process [19].

Taken together, these findings suggest that artificial intelligence-based virtual screening has obvious advantages in the discovery of active ingredients in TCM - higher efficiency and lower experimental costs.

3.3. Efficacy evaluation testing

Pharmacological evaluation test is a key stage in the research and development of new drugs of TCM, involving the evaluation of ADMET (absorption, distribution, metabolism, excretion and toxicity) properties [20]. Traditional ADMET evaluation requires a large number of animal experiments and clinical trials, which are costly, time-consuming and have great ethical controversy. The artificial intelligence-assisted ADMET prediction model provides an efficient and cost-effective solution for this stage.

An AI-based method was introduced to identify hepatotoxic compounds in TCM–western medicine combinations [21]. The hepatotoxicity prediction system was constructed by using chemical structure characteristics and AI model. This method provides a new technical approach for the safety evaluation of TCM.A molecular network similarity method based on deep learning was applied to rapidly screen illegal adulterants in dietary supplements and Chinese herbal medicines [22]. The role of AI in the entire research and development of TCM was examined in a review [23]. It identifies three key stages - drug design, drug manufacturing, and market access - each shaped by AI-driven innovation. Studies have shown that artificial intelligence technology can significantly improve the efficiency and success rate of TCM research and development, and has unique advantages in ADMET prediction and safety evaluation. The application of AI in computer-aided drug design was further confirmed to effectively supplement traditional methods, promoting the intelligent development of TCM [24].

In order to systematically present the methodological characteristics and practical value of AI technology in different development stages, this paper summarizes the main AI methodologies and their specific application scenarios in each stage, as shown in Table 1.

Table 1. Summarises the main applications of artificial intelligence in the three key stages of TCM R&D

Key Stage	AI Methodology	Main Applications
Target Identification and Characterisation	Machine learning, deep learning, graph neural networks, NLP	Multi-omics data integration, knowledge graph construction, target prediction
Screening of Active Compounds	Virtual screening, molecular docking, QSAR modelling	Compound library screening, bioactivity prediction, structure optimisation
Pharmacological evaluation testing	ADMET prediction models, toxicity prediction	Absorption and distribution prediction, metabolism and excretion analysis, safety evaluation

4. Discussion

Based on the preceding analysis, this paper explores the development directions, challenges and response strategies regarding the integration of artificial intelligence and TCM.

DOI: 10.54254/2753-8818/2026.AU35394

4.1. Development directions for the integration of AI and TCM

Traditional drug development has a big problem — it keeps staring at "one drug, one target, one disease." But real-world diseases aren't that simple. It just doesn't work.Chinese medicine, on the other hand, has always been multi-target by nature. For these complex diseases, it actually fits better. Now with network pharmacology and AI, you can map out those "component–target–pathway–disease" relationships. Slowly, it pulls research away from that single-target mindset and toward multi-target strategies.

TCM formulas have long used the "sovereign, minister, assistant, and messenger" structure. But honestly, that came from experience, not from lab-verified science. And that's been a big hurdle for TCM to get international recognition. So what can machine learning do? It can find statistical patterns in large formula datasets. That means you can design formulas in an algorithm-guided way — results become reproducible and scientifically defensible. That actually holds up. Another interesting link. TCM talks about "preventing disease before it occurs." Basically it's about keeping overall health in check. That lines up really well with modern preventive medicine.Multi-omics data, lifestyle information and environmental data are put into the AI model to obtain personalized health assessment. This has changed the role of TCM - it is no longer just to treat diseases. It can now be used as part of precision health management. The traditional concept has only been woven into a modern technology framework.

Big data and artificial intelligence have the potential to change drug discovery based on TCM [25]. These technologies make the research workflow more innovative and rational, with the broader goal of producing higher quality data throughout the development process. This expansion from disease treatment to health status intervention will open up new ways for the application of TCM in health management, chronic disease prevention and anti-aging.

4.2. Challenges and response strategies

Although artificial intelligence has shown great potential in the research and development of TCM, it still faces many challenges:

Data quality problem: TCM data has the characteristics of low standardization and inconsistent quality, which affects the training effect of AI model. The countermeasures include establishing a unified standard for TCM data, strengthening data quality control, and building a high-quality TCM knowledge map.

Model interpretability: Deep learning models are often regarded as ' black boxes ', and their decision-making processes are difficult to explain. Countermeasures include the development of interpretable artificial intelligence technology and the integration of TCM theory to enhance the transparency and credibility of the model.

Lack of compound talents: The research and development of artificial intelligence in TCM requires multidisciplinary professionals with both TCM and artificial intelligence expertise. The countermeasures include strengthening interdisciplinary education and training, and promoting the deep integration of TCM and computer science. computer science.

5. Conclusion

So this paper — it looks at how AI is helping with TCM R&D. We break it down by phase.First, target identification. AI speeds things up a lot. Finds targets faster, cuts down on lab costs.Then there's screening for active ingredients. AI does virtual screening, predicts which molecules might

DOI: 10.54254/2753-8818/2026.AU35394

actually work. That makes the whole process way more efficient. And for pharmacodynamic evaluation? AI-driven ADMET models — they can assess five different drug characters pretty well. That means better safety.So those are the main bits.

We also talk about where AI and TCM are heading together. Like shifting from single-target to multi-target strategies. Moving from empirical formulas to algorithm-based ones. And going beyond just treating disease — toward managing the body's state. These changes could really modernize TCM. And help China's pharma industry move from making generics to real innovative drugs.

Now, don't get me wrong — there are still challenges. Data quality. Model interpretability. Not enough people who know both AI and TCM. But tech keeps improving, and research keeps digging deeper.

Here's the thing: AI in TCM R&D has already moved past just understanding biological networks — you know, mapping them out. Now it's actually about actively controlling those networks. It's not just data analysis anymore — it's intelligent optimisation. That's a real paradigm shift.

References

[1] Huo, X., Gu, Y., & Zhang, Y. (2022, July 15). The discovery of multi-target compounds with anti-inflammatory activity from traditional Chinese medicine by TCM-target effects relationship spectrum. *Journal of Ethnopharmacology, 293*, 115289. https://doi.org/10.1016/j.jep.2022.115289

[2] Ren, F., Aliper, A., Chen, J., Zhao, H., Rao, S., Kuppe, C., et al. (2025). A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models. *Nature Biotechnology, 43*(1), 63–75.

[3] Lin, Y., Zhang, Y., Wang, D., Yang, B., & Shen, Y. Q. (2022, December). Computer-assisted, particularly AI-assisted, drug virtual screening and design in traditional Chinese medicine. *Phytomedicine, 107*, 154481.

[4] Bai, X., Ta, N., Gong, G. H., Zhang, B., & Wei, C. X. (2022). Effects of integrated traditional Chinese medicine and conventional Western medicine on the quality of life of breast cancer patients: A systematic review and meta-analysis. *Evidence-Based Complementary and Alternative Medicine, 2022*, 3123878.

[5] Jin, M., Ren, W., Zhang, W., Liu, L., Yin, Z., & Li, D. (2021, November 9). Exploring the underlying mechanism of Shenyankangfu Tablet in the treatment of glomerulonephritis through network pharmacology, machine learning, molecular docking, and experimental validation. *Drug Design, Development and Therapy, 15*, 4585–4601.

[6] Isozaki, A., Harmon, J., Zhou, Y., Li, S., Nakagawa, Y., Hayashi, M., Mikami, H., Lei, C., & Goda, K. (2020, August 26). AI on a chip. *Lab Chip, 20*(17), 3074–3090.

[7] Chakraborty, C., Bhattacharya, M., Lee, S. S., Wen, Z. H., & Lo, Y. H. (2024, August 8). The changing landscape of drug discovery using AI and deep learning: Recent advancements, success stories, collaborations, and challenges. *Mol Ther Nucleic Acids, 35*(3), 102295.

[8] Jiménez-Luna, J., Grisoni, F., Weskamp, N., & Schneider, G. (2021, September). Artificial intelligence in drug discovery: Recent advances and future perspectives. *Expert Opin Drug Discov, 16*(9), 949–959. https://doi.org/10.1080/17460441.2021.1909567

[9] Damiati, S. A. (2020, July 26). Digital pharmaceutical sciences. *AAPS PharmSciTech, 21*(6), 206.

[10] Gallego, V., Naveiro, R., Roca, C., Ríos Insua, D., & Campillo, N. E. (2021, August). AI in drug development: A multidisciplinary perspective. *Mol Divers, 25*(3), 1461–1479.

[11] Lam, C. S., Peng, L. W., Yang, L. S., Chou, H. W. J., Li, C. K., Zuo, Z., Koon, H. K., & Cheung, Y. T. (2022, Sep). Examining patterns of traditional Chinese medicine use in paediatric oncology: A systematic review, meta-analysis and data-mining study. *J Integr Med, 20*(5), 402-415.

[12] Li, Y., Liu, X., Zhou, J., Li, F., Wang, Y., & Liu, Q. (2025, April 15). Artificial intelligence in traditional Chinese medicine: Advances in multi-metabolite multi-target interaction modelling. *Front Pharmacol, 16*, 1541509.

[13] Liu, K., Li, Y., Lv, Y., et al. (2022). Multi-target-based polypharmacology prediction (mTPP): An approach using virtual screening and machine learning for multi-target drug discovery. *Chemico-Biological Interactions, 368*, 110239. doi: 10.1016/j.cbi.2022.110239

[14] Liu, Z., Luo, C., Fu, D., Gui, J., Zheng, Z., Qi, L., & Guo, H. (2022). A novel transfer learning model for traditional herbal medicine prescription generation from unstructured resources and knowledge. *Artificial Intelligence in Medicine, 124*, 102232. ISSN 0933-3657

[15] Ma, S., Liu, J., Li, W., et al. (2023). Machine learning in Traditional Chinese Medicine with natural products and molecules: Current status and future perspectives. *Chinese Medicine, 18*(1), 1-18.

DOI: 10.54254/2753-8818/2026.AU35394

[16] Kim, E., Choi, A., & Nam, H. (2019). Drug repositioning of herbal compounds via a machine-learning approach. *BMC Bioinformatics, 20*(1), 1–12.
[17] Zhang, H., Pan, J., Wu, X., et al. (2019). Large-scale target identification of herbal medicine using a reverse docking approach. *ACS Omega, 4*(4), 6710–6717.
[18] Erlina, Paramita, R. I., Kusuma, W. A., et al. (2022). Virtual screening of Indonesian herbal compounds as COVID-19 supportive therapy: Machine learning and pharmacophore modelling approaches. *BMC Complementary Medicine and Therapies, 22*(1), 1–15.
[19] Hioki, K., Hayashi, T., Natsume-Kitatani, Y., et al. (2022). Machine learning-assisted screening of herbal medicine extracts as vaccine adjuvants. *Frontiers in Immunology, 13*, 847616.
[20] Chen, W., Liu, X., Zhang, S., et al. (2023). Artificial intelligence for drug discovery: Resources, methods, and applications. *Molecular Therapy-Nucleic Acids, 33*, 1–15.
[21] Chen, Z., Zhao, M., You, L., et al. (2022). Developing an artificial intelligence method for screening hepatotoxic compounds in traditional Chinese medicine and Western medicine combinations. *Chinese Medicine, 17*(1), 1–14.
[22] Sheng, Y., Xue, Y., Wang, J., et al. (2023). Fast screening and identification of illegal adulteration in dietary supplements and herbal medicines using molecular networking with deep-learning-based similarity enhancement. *Analytical and Bioanalytical Chemistry, 415*(18), 1–14.
[23] Wang, A., Luo, Q., Tan, X., et al. (2026). Development and application of artificial intelligence in traditional Chinese medicine research and development. *Chinese Medicine, 21*(1), 1–15.
[24] He, X., Sun, M., Battulga, T., et al. (2025). Application of artificial intelligence in the development of traditional Chinese medicine. *Basic & Clinical Pharmacology & Toxicology, 136*(2), e70066.
[25] Li, M., & Zhang, J. (2023). A focus on harnessing big data and artificial intelligence: Revolutionising drug discovery from traditional Chinese medicine sources. *Chemical Science, 14*(15), 4011–4024.