10.70711/mhr.v2i3.4915

Research on the Application of Artificial Intelligence Technology in Drug Development

Douqiu Liu

Pharmaceutical College, Tianjin Medical University, Tianjin 300041, China

liudouqiu@tmu.edu.cn

Abstract: Artificial intelligence (AI) has made significant progress in drug research and development in recent years. This paper reviews relevant literature from the past five years and discusses the applications and future development trends of AI in the four major steps of drugs. The discovery of drugs, preclinical research, clinical trials, and drug approval and marketing.

It shows the innovations and efficiency improvements of AI techniques at each stage.

Keywords: Drug discovery; Artificial intelligence; Target identification; Deep learning; Toxicity prediction

1. Introduction

The process of drug research and development often has some bad features, such as a long cycle, high cost, and low success rates. However, with the development of information technology such as internet big data and cloud computing, AI has become a hot direction in this field. It is now being applied not only to multiple stages of drug discovery but also to more significant breakthroughs and advances in the treatment of disease. This paper examines the use of AI, which is currently popular in drug research and development, and summarizes its main accomplishments and research progress.

2. Theoretical Research on Artificial Intelligence Technology and Drug Development

2.1 Concept and characteristics of artificial intelligence technology

AI is a new technical science that studies and develops theories, methods, technologies, and application systems used to simulate, extend, and expand human intelligence. Its research includes a wide range of areas, like search algorithms, knowledge graphs, natural language processing, expert systems, evolutionary algorithms, machine learning, deep learning, reinforcement learning, molecular dynamics simulations, and big data analysis.

2.2 Drug discovery process

The process of drug discovery is complex and lengthy. It involves four main steps: (1) Drug Discovery: Which mainly includes identifying disease-related genes, validating potential targets, and discovering and refining lead compounds. (2) Preclinical Studies: Evaluating the safety and efficacy of drug candidates through in vitro (test tube or cell culture) and in vivo (animal) tests. (3) Clinical Trials: These are divided into three phases (I, II, III) and involve assessing the safety and efficacy of drug candidates in different human groups. (4) Drug Approval and Marketing: Agencies such as the FDA review safe and effective drugs for approval and commercialization while continuing to monitor their post-marketing status.

According to data, this process typically needs decades and billions of dollars to invest. Only 10% of drugs successfully pass the Phase I clinical trial. Many of them fail due to high toxicity or low efficacy. And for complex diseases like cancer, the success rate drops to just 5%. The traditional drug development model heavily relies on the personal experience and creativity of developers. It always leads to long times, high costs, low efficiency, and high risk. Thus, finding new technologies and methods to reduce costs and accelerate drug development is a major priority for pharmaceutical companies.

2.3 Current status

The integration of AI with data, algorithms, and computational power is gradually transforming drug research and development. Nearly every major pharmaceutical company worldwide has forged close partnerships with AI firms. For example, AstraZeneca published 65 AI-focused papers on drug R&D in 2019 alone. In 2020, Google DeepMind's AlphaFold2 achieved a major breakthrough by solving the complex problem of protein folding, which had stumped scientists for over fifty years [1]. AI technologies are expected to greatly speed up the drug dis-

covery process, reduce costs, and improve success rates by increasing the speed and accuracy of data analysis. Its ultimate goal is to facilitate the rapid development and scale-up of new drugs through innovative approaches.

3. Applications of Artificial Intelligence in Drug Discovery

3.1 Drug discovery

Screening and validating potential drug targets are the initial steps in drug discovery. This involves identifying the target, characterizing the protein's structure, and developing drug molecules to modulate its function. The success of future research depends heavily on these foundational steps. Traditional methods rely more on extensive experimentation and literature reviews. They are often time-consuming, costly, and inefficient. By combining systems biology with AI algorithms, we can dig deeper into the correlation between multi-omics data and patient clinical health information and use natural language processing methods to analyze uncovered pathways, proteins, and action mechanisms, to find new action mechanisms and targets, to improve the efficiency and accuracy of treatment.

3.1.1 Based on Systems Biology

Systems biology is the study of the relationship between the components of the living system. It aims to create a comprehensive model and an overall understanding of the organism. Interrogative Biology, developed by BERG, an AI biotech company, is an AI-based systems biology platform that generates data-driven, non-biased networks to identify targets and disease markers [2]. Additionally, Benevolent AI's judgment-augmented cognition system (JACS) helps with drug repositioning by uncovering new connections within extensive unstructured data, including diseases, drugs, and trial information [3]. In 2020, the researchers used this AI platform to find that the classic JAK kinase inhibitor baricitinib could be used to treat novel coronavirus infections [4].

3.1.2 Based on Deep Learning

At the heart of AI is machine learning, which can be divided into supervised and unsupervised learning. Deep learning is a supervised learning method, that learns the internal rules and expressions of sample data to effectively understand text, images, sounds, and other information. So that machines can have similar analysis and learning capabilities as humans and achieve effective identification of a variety of data.

For example, Wan et al [5] combined efficient feature embedding and deep neural networks with a new scalable computing framework, Deep CPI, and took three G-protein-coupled receptors as research objects to achieve efficient identification of drug targets through bioinformatics analysis. Additionally, Liu et al [6] used two deep learning models, Chemprop and RTMScore, to evaluate the activity of SMYD2 inhibitors based on the interaction between small molecules, and conducted in vitro cell proliferation and migration experiments on small molecules with good activity, to provide new ideas for the research and development of novel anti-lung cancer drugs.

3.1.3 Based on Proteins

Proteins have complex spatial structure, which determines their different biological functions. Therefore, it is of great significance for the research and development of new drugs to accurately obtain the three-dimensional spatial structure of target proteins.

Following almost fifty years of research and development, experimental methods such as cryo-electron microscopy (Cryo-EM), X-ray crystallography, and nuclear magnetic resonance (NMR) have been used to investigate the structure of proteins. These techniques are insufficient, though. They are expensive and time-consuming. Additionally, they are ill-equipped to handle people's deep comprehension of both sophisticated drug discoveries and life processes. In recent studies, the AI tool AlphaFold successfully predicted the 3D structures of drug target proteins [7]. Published in the journal Nature on May 8, 2024, AlphaFold3, developed by a joint team from Google Deep Mind and the London-based drug discovery company Isomorphic Labs. It can not only predict the structure of proteins, but also accurately forecast the functions and interactions of all life's molecules, including nucleic acids, small molecules, ions, modified protein residue complexes, and antibody antigens.

AI has shown significant potential and widespread opportunities in the identification and validation of drug targets. Through the fusion and analysis of massive biomedical big data, AI technologies not only improve the efficiency and accuracy of target identification but also provide new prospects and challenges for new drug development. As technology advances, the use of AI in this field is set to become even more comprehensive and impactful.

3.2 Preclinical studies

Preclinical research is an important phase in the development of drugs. It includes pharmacokinetic (PK) and pharmacodynamic (PD) studies to assess the safety and effectiveness of potential drug candidates. By simulating and predicting, the AI technique can greatly increase clinical efficiency, lower the experimental cost, and provide more precise results.

3.2.1 Pharmacokinetic simulation

The ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) pharmacokinetic method is essential in drug design and screening. AI significantly enhances ADMET prediction by using big data and machine learning. It creates various models for analyzing and

forecasting these properties. This support is vital for decision-making in drug design and development. Therefore, the use of AI technology to accelerate and improve ADMET forecasting has become an attractive trend.

Chou et al proposed a new research paradigm that combines the physiological pharmacokinetic (PBPK) model with machine learning or AI [8]. This combination can accurately predict and improve the accuracy of in-vivo absorption, distribution, metabolism, excretion, and other characteristics of drug candidates in a relatively short time. It is helpful to find compounds with optimal pharmacokinetic properties as early as possible and accelerate the process of new drug research and development. Additionally, Wang et al developed a new metabolite prediction method based on deep learning [9]. This algorithm can generate a wide range of metabolites and use deep neural networks to sort them. Finally, it can obtain the optimal metabolites, providing a theoretical basis for further optimization of their metabolic characteristics.

3.2.2 Pharmacodynamic studies

Lind et al combined cancer screening data with machine learning to build a random forest model that can predict drug activity based on the genomic variation of tumor cells, providing a convenient method for personalized cancer treatment, drug reuse, and drug research and development [10]. Additionally, Mayr et al developed a DeepTox process based on deep learning and predicted 12 toxic effects of 12 compounds and drugs in the Tox21 database Challenge, achieving the best results, exceeding naive Bayes, SVM, random forest, and other algorithms. The DeepTox algorithm has high accuracy in predicting drug toxicology, reducing toxicity problems in clinical trials, and improving screening efficiency in the preclinical stage [11].

Furthermore, the use of AI for research on pharmaceutical properties, toxicology, and other aspects can also reduce the amount of experimental animals. For example, Insilico Medicine company's Life Star can automatically train high-quality biological information in the laboratory into AI, and conduct animal tests. It can accurately predict the efficacy and toxic side effects of new drugs, reducing the dependence on animal experiments and the cost and time of new drug research and development.

AI plays a crucial role in preclinical research. Through simulation and prediction, it can quickly assess the safety and efficacy of drug candidates, reducing experimental investment and accelerating the drug development process.

3.3 Clinical trials

In the process of new drug development, clinical trials are the most important and costly link. Traditional experimental design and implementation have problems such as low efficiency and low success rate. AI technology can greatly improve the efficiency and success rate of clinical trials through intelligent design, patient recruitment, data analysis, and other means.

In 2021, Jenson and Komodo Health agreed on a new research program that will integrate AI algorithms with patient data into clinical trials to improve the effectiveness and success of studies. In addition, Tempus uses predictive algorithms to find biomarkers to help select suitable patients, especially for oncology research [12].

Deep 6 AI utilizes natural language processing (NLP) and machine learning algorithms to search for patients who meet the criteria for clinical trials. Such an approach significantly improves the speed and accuracy of the admission of patients and ultimately reduces the time and costs involved in clinical trials [13]. In addition, Medidata Solutions uses AI to continuously monitor and ensure the quality of real-time data, and automatically detect and correct data errors to improve the accuracy and completeness of data. Such real-time data management has significantly improved the effectiveness and quality of clinical studies [14].

AI has been widely applied in clinical trials. It can optimize the design of clinical trials, increase the efficiency of admission, monitor the outcome, increase the success rate, lower the cost of the trial, and guarantee the success of the medicine.

3.4 Drug Approval and Marketing

The approval and marketing of drugs is the final stage of new drug development. It involves a series of regulatory and approval processes. The application of AI technology is mainly focused on data analysis, documentation regulatory compliance inspection, and other aspects. It can speed up the approval process and improve the efficiency and compliance of supervision.

Open Targets, for example, uses NLP to analyze clinical trial failures and then fuses the data into the platform to further improve clinical trial design. Johnson & Johnson uses AI automatic processing technology to classify and analyze clinical trial data and form reports and documents submitted to regulatory authorities. It greatly shortens the time for data summary and report writing, speeding up the drug approval process [13]. In addition, Pfizer also uses machine learning algorithms to analyze clinical trial data to verify that it meets relevant regulations such as the FDA, thereby improving the authenticity of the data and reducing the approval delays caused by data discrepancies [15].

AI technology has played a huge role in the whole process of drug development, greatly speeding up the entire process of drugs from research to marketing. The above research results indicate that AI technology has broad application prospects, and also provides a strong support for future new drug research and development.

4. Conclusion

The application of AI in drug development is increasingly widespread, covering the entire process from target screening to drug marketing. Through the introduction of AI technology, the efficiency and success rate of drug research and development have been significantly improved, and the cost has been significantly reduced. While AI shows great potential in drug discovery, it still faces technical and ethical challenges. Insufficient data quality and quantity, algorithmic transparency and interpretability issues, and data privacy issues are all challenges that need to be addressed. In the future, with the continuous development and improvement of AI technology, its application in drug research and development will be more in-depth and extensive. The drug research and development process will be more efficient and accurate, promoting the innovation and development of the pharmaceutical industry, which is worth further exploration and research.

References

[1] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. *Nature*. 2021;596(7873):583-589.

[2] Narain N, Kiebish M, Vishnudas V, et al. CSAO-1. Interrogative Biology: Unraveling insights into causal disease drivers by use of a dynamic systems biology and Bayesian AI to identify the intersect of disease and healthy signatures[J]. *Neuro-Oncology Advances*. 2021;3(Supplement_2):ii1.

[3] Wang C, Xiao F, Li MZ, et al. Progress in the application of artificial intelligence in screening and validation of drug targets[J]. *J China Pharm Univ.* 2023;54(3):269-281.

[4] Richardson P, Griffin I, Tucker C, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease[J]. *Lancet*. 2020;395(10223):e30-e31.

[5] Wan F, Zhu Y, Hu H, et al. DeepCPI: A Deep Learning-based Framework for Large-scale in silico Drug Screening[J]. *Genomics Proteomics Bioinformatics*. 2019;17(5):478-495.

[6] Liu XQ, Zhu YJ, Feng DW, et al. Screening and activity evaluation of SMYD2 inhibitors based on deep learning[J]. *Journal of Yantai University (Natural Science and Engineering Edition).*:1-11.

[7] Ding BX, Hu J, Wang JF. Progress in the application of artificial intelligence in drug discovery and development[J]. *Shandong Chemical Industry.* 2019;48(22):70-73.

[8] Chou WC, Lin Z. Machine learning and artificial intelligence in physiologically based pharmacokinetic modeling[J]. *Toxicol Sci*. 2023;191(1):1-14.

[9] Wang D, Liu W, Shen Z, et al. Deep Learning Based Drug Metabolites Prediction[J]. *Front Pharmacol*. 2019;10:1586.

[10] Lind AP, Anderson PC. Predicting drug activity against cancer cells by random forest models based on minimal genomic information and chemical properties[J]. *PLoS One*. 2019;14(7):e0219774.

[11] Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: Toxicity Prediction using Deep Learning[J]. *Front Environ Sci*. 2016;3.

[12] Savage N. Tapping into the drug discovery potential of AI[J]. *Biopharma Dealmakers*. Published online May 27, 2021.

[13] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence[J]. *Nat Med*. 2019;25(1):44-56.

[14] Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare[J]. *Nat Med*. 2019;25(1):24-29.

[15] Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects[J]. *Drug Discov Today*. 2019;24(3):773-780.