Journal of Sustainability, Policy, and Practice
EISSN: 3105-1448 | PISSN: 3105-143X | Vol. 2, No. 1 (2026)

Review

Applications of Artificial Intelligence Techniques in Medical Drug Discovery and Precision Treatment

Xiaojing Li [1,*]

[1] School of Biological Sciences, Universiti Sains Malaysia, Penang, Malaysia
* Correspondence: Xiaojing Li, School of Biological Sciences, Universiti Sains Malaysia, Penang, Malaysia

Abstract: Artificial intelligence (AI) has revolutionized various facets of medical drug discovery and precision treatment. This review paper provides a comprehensive overview of AI techniques applied in these domains, encompassing historical developments, core applications, comparative analyses, existing challenges, and future perspectives. The review begins with a historical overview of AI in medicine, charting its evolution from expert systems to deep learning. Core applications are then explored, including AI-driven drug target identification, de novo drug design, prediction of drug efficacy and toxicity, and patient stratification for precision treatment. Specific AI techniques such as machine learning, deep learning, and natural language processing are examined in the context of each application. A critical comparison of different AI approaches highlights their strengths and limitations. The review also addresses challenges in the field, such as data biases, lack of interpretability, and regulatory hurdles. Finally, future directions are discussed, emphasizing the potential of AI to transform drug discovery and personalized medicine. This review aims to serve as essential reference for researchers and practitioners in the intersection of AI and medicine, inspiring future advancements in the field.

Keywords: artificial intelligence; drug discovery; precision treatment; machine learning; deep learning; personalized medicine; AI in medicine

Received: 03 December 2025
Revised: 12 January 2026
Accepted: 26 January 2026
Published: 30 January 2026

1. Introduction

1.1. Background and Motivation

The intersection of artificial intelligence (AI) and medicine holds immense promise, particularly in revolutionizing drug discovery and enabling precision treatment strategies. Traditional drug development is a lengthy, expensive, and often inefficient process, characterized by high failure rates and significant resource investment. Similarly, conventional treatment approaches often follow a one-size-fits-all model, neglecting the individual variability in patient responses [1]. AI offers the potential to overcome these limitations by accelerating the identification of drug candidates, predicting treatment outcomes, and tailoring therapies based on individual patient profiles, considering factors like genetics (g), lifestyle (l), and environment (e).

1.2. Scope and Objectives

This review focuses on the application of artificial intelligence (AI) techniques, specifically machine learning, deep learning, and natural language processing, in medical drug discovery and precision treatment. We will cover applications including target identification, drug design, clinical trial optimization, and personalized medicine. The objectives are to provide a comprehensive overview of current AI applications in these

areas, identify key challenges hindering widespread adoption, and suggest future research directions to accelerate AI-driven advancements in healthcare, ultimately improving patient outcomes and reducing healthcare costs, where $cost$ is a function of $time$ and $resources$.

2. Historical Overview of AI in Medicine

2.1. Early Expert Systems

Early AI in medicine centered on expert systems designed for diagnosis and treatment. MYCIN, for example, aimed to diagnose bacterial infections and recommend antibiotics. These systems encoded expert knowledge into rules, enabling automated reasoning (see Table 1). Successes included demonstrating the feasibility of AI-driven decision support. However, limitations arose from knowledge acquisition bottlenecks, difficulty handling uncertainty, and limited generalizability beyond narrow domains. The $P(D \mid S)$probability of disease D given symptom S was often hard to quantify.

Table 1. Evolution of AI Technologies in Medicine.

Era	Technology	Description	Key Features	Successes	Limitations	Challenges
Early AI	Expert Systems (e.g., MYCIN)	Rule-based systems for diagnosis and treatment recommendations.	Encoding expert knowledge into rules for automated reasoning.	Demonstrated feasibility of AI-driven decision support.	Knowledge acquisition bottlenecks, difficulty handling uncertainty, limited generalizability.	Quantifying probabilities like $P(D\|S)$(probability of disease D given symptom S).

2.2. Machine Learning Era

The machine learning era marked a significant paradigm shift, moving from rule-based expert systems to data-driven approaches. Statistical models, such as logistic regression for classification and linear regression for prediction of drug efficacy ($y = mx + b$), gained prominence. Machine learning offered advantages like adaptability to new data and the ability to uncover hidden patterns. However, it also presented challenges, including the need for large datasets and the "black box" nature of some algorithms, contrasting with the transparency of expert systems [2].

2.3. Deep Learning Revolution

The deep learning revolution significantly advanced medical AI. These models excel at learning intricate patterns from massive datasets, surpassing traditional machine learning. This capability is crucial for drug discovery, enabling prediction of drug-target interactions and molecular properties with unprecedented accuracy. In precision treatment, deep learning analyzes patient data (x_i), like genomics and imaging, to personalize treatment plans and predict therapeutic response ($y_i = f(x_i)$), paving the way for more effective and targeted therapies.

3. AI-Driven Drug Target Identification and Validation

3.1. Target Identification using Genomics and Proteomics Data

AI algorithms are revolutionizing drug target identification by leveraging the vast amounts of genomic and proteomic data available. Network analysis identifies key proteins within biological pathways, where nodes represent proteins and edges represent interactions. Differential expression analysis pinpoints genes and proteins that are significantly upregulated or downregulated in disease states compared to healthy

controls, calculated using statistical tests like t-tests or ANOVA, with $p<0.05$ often considered significant (see Table 2) [3]. Machine learning models, including support vector machines (SVMs) and deep neural networks (DNNs), are trained on multi-omics data to predict novel drug targets. These models learn complex relationships between genomic features, protein expression levels, and disease phenotypes, assigning a probability score P (target | data) indicating the likelihood of a protein being a viable target. This integrated approach accelerates the target discovery process.

Table 2. Comparison of Genomics and Proteomics in Target Identification with AI.

Feature	Genomics	Proteomics
Data Type	DNA sequences, gene expression levels (mRNA)	Protein sequences, protein expression levels
Information Provided	Genetic predisposition to disease, potential for gene-based therapies	Actual protein activity and abundance, direct indicators of cellular function
Role in Target Identification	Identifies genes that are differentially expressed or mutated in disease	Quantifies protein expression changes, identifies post-translational modifications
AI Application	Predicts gene targets based on sequence analysis and gene expression profiles	Predicts protein targets based on protein expression patterns and interactions.
Network Analysis	Identifies key genes within gene regulatory networks.	Identifies key proteins within protein-protein interaction networks.
Differential Expression Analysis	Detects genes with significantly altered expression levels in disease states, calculated using statistical tests like t-tests or ANOVA, with $p < 0.05$ often considered significant.	Detects proteins with significantly altered expression levels in disease states, calculated using statistical tests like t-tests or ANOVA, with $p < 0.05$ often considered significant.
Machine Learning Models	Trained on genomic datasets to predict drug target potential and assign a probability score $P(target\|data)$.	Trained on proteomic datasets to predict drug target potential and assign a probability score $P(target\|data)$.
Limitations	Does not always reflect protein levels or activity due to post-transcriptional regulation.	Technically challenging to quantify all proteins comprehensively; may miss rare or transient protein modifications.

3.2. Virtual Screening and Hit Discovery

AI significantly accelerates virtual screening for hit discovery by computationally evaluating large compound libraries against a target protein. Docking simulations predict the binding pose and affinity of each compound to the target's binding site. These simulations utilize scoring functions, which are mathematical models estimating the binding free energy (ΔG_{bind}) based on factors like shape complementarity, hydrogen bonding, and hydrophobic interactions. Traditional scoring functions are often enhanced by machine learning (ML) models. ML algorithms, trained on experimental binding data, can learn complex relationships between molecular features and binding affinity, leading to more accurate hit prediction [4]. These ML-based scoring functions can prioritize compounds with a higher probability of binding, significantly reducing the number of compounds requiring experimental validation and accelerating the drug discovery process.

3.3. De Novo Drug Design

AI plays a crucial role in de novo drug design, enabling the generation of novel drug candidates from scratch, tailored to specific targets and desired properties. Generative models, such as variational autoencoders (VAEs) and generative adversarial networks (GANs), learn the underlying distribution of chemical space from existing drug-like molecules and then sample new molecules with optimized characteristics like binding affinity and ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties. Reinforcement learning (RL) algorithms further refine these generated molecules by rewarding those that exhibit improved target binding or other desirable attributes. The RL agent learns to navigate the chemical space and optimize molecular structures based on feedback from a reward function. Evolutionary algorithms, inspired by natural selection, iteratively improve a population of molecules through processes of mutation and selection, driving the population towards molecules with higher fitness scores, often defined by their predicted activity against the target. These AI-driven approaches significantly accelerate the drug discovery process by exploring a vast chemical space and identifying promising drug candidates that might be missed by traditional methods [5].

4. AI in Predictive Modeling of Drug Efficacy and Toxicity

4.1. Predicting Drug Efficacy using Clinical Trial Data

AI algorithms are increasingly employed to analyze clinical trial data and predict drug efficacy, accelerating drug development and enabling precision treatment strategies. Survival analysis, often utilizing models like Cox proportional hazards, can predict time-to-event outcomes such as disease progression or death, allowing for the assessment of a drug's impact on patient survival. Regression models, including linear and logistic regression, are used to predict continuous or categorical efficacy endpoints based on patient characteristics and treatment regimens. Machine learning techniques, such as support vector machines and neural networks, offer more sophisticated approaches to response prediction. These algorithms can identify complex relationships between patient features (x_i), drug properties (d_j), and treatment outcomes (y), enabling the identification of patient subgroups that are most likely to respond favorably to a specific treatment. By uncovering these predictive biomarkers, AI facilitates the design of more targeted and effective clinical trials and personalized treatment plans [6].

4.2. Predicting Drug Toxicity using Preclinical Data

Predicting drug toxicity early in the development pipeline is crucial for reducing attrition rates and ensuring patient safety. AI techniques leverage preclinical data, including in vitro and in vivo studies, to build predictive models (see Table 3). Quantitative structure-activity relationship (QSAR) modeling correlates chemical structure with toxicity endpoints using statistical methods. Machine learning algorithms, such as support vector machines and random forests, can be trained on large datasets of chemical structures and toxicity data to predict the toxicity of new compounds. These models often use molecular descriptors like logP and molecular weight (MW) as input features. Deep learning approaches, particularly convolutional neural networks and recurrent neural networks, are increasingly used to automatically extract complex features from chemical structures and predict toxicity with high accuracy. These methods can identify subtle patterns in preclinical data that may be missed by traditional approaches, improving the reliability of toxicity predictions.

Table 3. Examples of AI-based toxicity prediction methods.

Method	Description	Input Features	Advantages
Quantitative Structure-Activity Relationship (QSAR)	Correlates chemical structure with	Molecular descriptors like log$P$ and	Relatively simple and interpretable.

	toxicity using statistical methods.	molecular weight (MW).	
Support Vector Machines (SVM)	Machine learning algorithm trained on chemical structures and toxicity data.	Chemical structures and toxicity data encoded as features.	Effective in high-dimensional spaces.
Random Forests	Ensemble learning method trained on chemical structures and toxicity data.	Chemical structures and toxicity data encoded as features.	Robust and less prone to overfitting.
Convolutional Neural Networks (CNNs)	Deep learning approach for automatic feature extraction from chemical structures.	Chemical structure representations (e.g., SMILES strings, molecular graphs).	Can automatically learn complex features from data.
Recurrent Neural Networks (RNNs)	Deep learning approach for automatic feature extraction from chemical structures.	Sequential representations of chemical structures (e.g., SMILES strings).	Effective in capturing sequential dependencies in molecular structures.

4.3. Integrating Multi-omics Data for Enhanced Prediction

Integrating multi-omics data, such as genomics, transcriptomics, proteomics, and metabolomics, offers a powerful approach to enhance the predictive accuracy of drug efficacy and toxicity models. By capturing a more holistic view of the biological system, these data types provide complementary information that can reveal complex drug-target interactions and downstream effects. AI algorithms, particularly machine learning models, are well-suited to handle the high dimensionality and heterogeneity of multi-omics datasets [7]. However, several challenges exist. These include data integration complexities, batch effects, and the need for sophisticated feature selection techniques to identify relevant biomarkers. Opportunities lie in developing novel AI architectures capable of effectively fusing multi-omics data, improving model interpretability, and ultimately accelerating drug discovery and personalized treatment strategies. The use of techniques like network analysis can further refine predictions by considering the interconnectedness of n biological entities represented by x_i, where $i = 1,2,\ldots,n$.

5. AI for Patient Stratification and Precision Treatment

5.1. Identifying Patient Subgroups using Machine Learning

Machine learning excels at identifying patient subgroups with varying treatment responses (see Table 4) [8]. Clustering algorithms group patients based on similarities in clinical data (X_i). Classification models predict treatment response (y) for individual patients. Dimensionality reduction techniques, such as PCA, handle high-dimensional data, improving model performance and revealing key features (f_i) driving subgroup differences. These methods enable targeted therapies [9].

Table 4. AI Methodologies Used in Patient Stratification.

Methodology	Description	Key Features/Data Used	Outcome
Clustering Algorithms	Group patients into subgroups based on similarities in clinical data.	Clinical data (X_i)	Identification of patient subgroups with distinct characteristics.

Classification Models	Predict treatment response (y) for individual patients.	Clinical data (X_i), Patient characteristics	Prediction of treatment response for individual patients.
Dimensionality Reduction (e.g., PCA)	Reduce the number of variables while preserving important information, improving model performance and revealing key features.	High-dimensional clinical data	Identification of key features (f_i) driving subgroup differences, improved model performance.

5.2. Predicting Treatment Response using Clinical and Molecular Data

AI algorithms predict treatment response by analyzing clinical (X) and molecular data (Y). Regression models estimate continuous responses, while classification models predict response categories (e.g., responder/non-responder). Deep learning, utilizing neural networks, captures complex relationships between multi-omics data and treatment outcomes. These models enhance precision treatment by identifying patients most likely to benefit from specific therapies [10].

5.3. AI-Driven Treatment Optimization

AI optimizes treatment by tailoring strategies to individual patient needs. Reinforcement learning algorithms can determine optimal dynamic treatment regimes, adjusting interventions based on patient response [11]. Personalized dosing, informed by AI analysis of patient-specific data like genetics and PK/PD parameters, ensures optimal drug efficacy while minimizing adverse effects. This leads to more effective and safer treatments [12].

6. Challenges and Limitations

6.1. Data Bias and Generalizability

Data bias presents a significant hurdle in applying AI to drug discovery and precision treatment [13]. Medical datasets often reflect existing health disparities, leading to skewed representations of certain demographics. This bias can manifest in various forms, including under-representation of specific ethnic groups, age ranges, or socioeconomic statuses. Consequently, AI models trained on such biased data may exhibit poor generalizability, performing accurately on the majority group while failing to provide reliable predictions for under-represented populations [14]. Addressing this requires concerted efforts to curate diverse and representative datasets that accurately reflect the heterogeneity of the patient population, ensuring equitable and effective AI-driven healthcare solutions for all. The impact of bias can be quantified using metrics like F_1 score disparity across different groups [15].

6.2. Interpretability and Explainability

The "black box" nature of many AI models poses a significant challenge in medical applications. Clinicians require transparent and understandable reasoning to trust AI-driven predictions and treatment recommendations. Interpretability ensures that the model's decision-making process is comprehensible, while explainability provides insights into *why* a particular prediction was made [16]. Techniques like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) can help illuminate feature importance and model behavior. Furthermore, inherently interpretable models, such as decision trees or linear models with regularization (e.g., L1 regularization where the penalty is $\lambda \sum_{i=1}^{n} |w_i|$ for weights w_i and regularization parameter λ), offer greater transparency from the outset [17].

7. Future Perspectives and Conclusion

7.1. Future Directions

AI's future in drug discovery and precision treatment lies in multi-omics data integration, generative AI for novel drug design, and personalized treatment optimization using reinforcement learning. Breakthroughs may arise from explainable AI, enhancing trust and adoption, and federated learning, enabling collaborative research while preserving data privacy.

7.2. Conclusion

AI offers immense potential in revolutionizing drug discovery and precision treatment. This review highlights AI's capacity to accelerate $R\&D$, improve diagnostic accuracy, and personalize treatment strategies. Future advancements promise even greater efficiency and efficacy in healthcare.

References

1. J. Deng, Z. Yang, I. Ojima, D. Samaras, and F. Wang, "Artificial intelligence in drug discovery: applications and techniques," Briefings Bioinform., vol. 23, no. 1, pp. 2022.
2. C. L. Cheong, "Research on AI security strategies and practical approaches for risk management," Journal of Computer, Signal, and System Research, vol. 2, no. 7, pp. 98–115, 2025.
3. J. Jin, T. Zhu, and C. Li, "Graph Neural Network-Based Prediction Framework for Protein-Ligand Binding Affinity: A Case Study on Pediatric Gastrointestinal Disease Targets," Journal of Medicine and Life Sciences, vol. 1, no. 3, pp. 136–142, 2025.
4. S. Li, K. Liu, and X. Chen, "A context-aware personalized recommendation framework integrating user clustering and BERT-based sentiment analysis," Journal of Computer, Signal, and System Research, vol. 2, no. 6, pp. 100–108, 2025.
5. H. S. Chan, H. Shan, T. Dahoun, H. Vogel, and S. Yuan, "Advancing drug discovery via artificial intelligence," Trends Pharmacol. Sci., vol. 40, no. 8, pp. 592-604, 2019.
6. V. Patel and M. Shah, "Artificial intelligence and machine learning in drug discovery and development," Intell. Med., vol. 2, no. 3, pp. 134-140, 2022.
7. R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, and P. Kumar, "Artificial intelligence to deep learning: machine intelligence approach for drug discovery," Mol. Divers., vol. 25, no. 3, pp. 1315-1360, 2021.
8. M. A. Sellwood, M. Ahmed, M. H. Segler, and N. Brown, "Artificial intelligence in drug discovery," Future Med. Chem., vol. 10, no. 17, pp. 2025-2028, 2018.
9. J. Jiménez-Luna, F. Grisoni, and G. Schneider, "Drug discovery with explainable artificial intelligence," Nat. Mach. Intell., vol. 2, no. 10, pp. 573-584, 2020.
10. K. K. Mak, Y. H. Wong, and M. R. Pichika, "Artificial intelligence in drug discovery and development," Drug Discovery and Evaluation: Safety and Pharmacokinetic Assays, pp. 1461-1498, 2024.
11. J. L. Zhao, "Graph-based deep dive on AI startup revenue composition and venture capital network effect," Economics and Management Innovation, vol. 3, no. 1, pp. 27–36, 2026.
12. C. Sarkar, B. Das, V. S. Rawat, J. B. Wahlang, A. Nongpiur, I. Tiewsoh, et al., "Artificial intelligence and machine learning technology driven modern drug discovery and development," Int. J. Mol. Sci., vol. 24, no. 3, 2023.
13. Y. Jing, Y. Bian, Z. Hu, L. Wang, and X. Q. S. Xie, "Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era," AAPS J., vol. 20, no. 3, 2018.
14. W. Chen, X. Liu, S. Zhang, and S. Chen, "Artificial intelligence for drug discovery: Resources, methods, and applications," Mol. Ther. Nucleic Acids, vol. 31, pp. 691-702, 2023.
15. C. Hasselgren and T. I. Oprea, "Artificial intelligence for drug discovery: are we there yet?," Annu. Rev. Pharmacol. Toxicol., vol. 64, no. 1, pp. 527-550, 2024.
16. R. Luo, X. Chen, and Z. Ding, "SeqUDA-Rec: Sequential user behavior enhanced recommendation via global unsupervised data augmentation for personalized content marketing," arXiv preprint arXiv:2509.17361, 2025.
17. Y. S. Cai, "Organizational restructuring of fintech enterprises: A strategic study balancing compliance and innovation," Financial Economics Insights, vol. 3, no. 1, pp. 1–10, 2026.