INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS

Review Paper

Applications of AI in Pharmaceuticals and Drug Discovery: A Comprehensive Review

C. Sadak Vali*[1], Abdullah Khan[2], S. Siva Prasad[3], Mare Pratibha Bharathi[4], Reddy Nazemoon[5], Survi Swati Goud[6]

[1] *Principal & Professor, Madhav Rao Patil College of Pharmacy, Maharashtra, India.*
[2] *HOD of Pharmaceutics, KPJ Healthcare University College, Nilai, Malaysia.*
[3] *Professor, MNR College of Pharmacy, Hyderabad. Telangana, India.*
[4] *Professor, Gokaraju Gangaraju College of Pharmacy, Hyderabad, Telangana, India.*
[5] *Professor & HOD, Bharat Institute of Pharmacy, Hyderabad, Telangana, India.*
[6] *Assistant Professor, UCPS, JNTU (Sultanpur), India.*

ARTICLE INFO

Published: 19 Nov 2025
Keywords:
Artificial Intelligence; Machine Learning; Deep Learning; Drug Discovery; Target Identification; Lead Optimization; Drug Repurposing; Pharmacokinetics; Clinical Trials.
DOI:
10.5281/zenodo.17649085

ABSTRACT

The drug discovery and development process is notoriously expensive, time-consuming, and prone to high failure rates. Artificial Intelligence (AI), particularly Machine Learning (ML) and Deep Learning (DL), is emerging as a transformative technology to address these persistent challenges. By rapidly processing vast and complex biomedical datasets, AI systems are augmenting human capabilities across the entire pharmaceutical pipeline. This review provides a comprehensive overview of the current applications of AI, specifically focusing on its impact on target identification, lead compound optimization, drug repurposing, and preclinical safety assessment. Furthermore, it discusses the integration of AI into clinical trial design and pharmaceutical manufacturing. While AI promises accelerated development timelines and reduced costs, its effective implementation is contingent upon overcoming significant hurdles related to data quality, model interpretability, and regulatory clarity. This paper concludes by highlighting the future directions and ethical- considerations essential for realizing the full potential of AI in creating safer, more efficacious, and accessible medicines.

INTRODUCTION

The journey of a novel therapeutic from conceptual idea to market approval typically spans over a decade and costs billions of U.S. dollars [1]. A major bottleneck is the high attrition rate, as approximately 90% of drug candidates fail during the preclinical or clinical development phases,

***Corresponding Author:** C. Sadak Vali
Address: *Principal & Professor, Madhav Rao Patil College of Pharmacy, Maharashtra, India.*
Email ✉: sadak2020@gmail.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

often due to lack of efficacy or unforeseen toxicity [2]. This inherent inefficiency of traditional discovery methods necessitates a paradigm shift.

Artificial Intelligence (AI), defined as a collection of computational tools that enable systems to perform tasks requiring human-like intelligence, is providing this necessary disruption. The exponential growth in biological data (genomic, proteomic, clinical, and chemical) has created the ideal environment for AI techniques, such as Machine Learning (ML) and Deep Learning (DL), to excel at pattern recognition, prediction, and generation [3]. AI systems can analyze heterogeneous data sources—from genetic information and protein structures to electronic health records—to uncover subtle relationships that are undetectable by conventional statistical or human analysis.

This review will systematically explore the various applications of AI across the pharmaceutical value chain, from the earliest stages of basic research to clinical development and beyond, concluding with a discussion of the current challenges and future outlook.

2. APPLICATIONS OF AI IN THE DRUG DISCOVERY PIPELINE:

The pharmaceutical pipeline can be broadly segmented into four phases where AI offers distinct

Advantages: Target Identification and Validation, Lead Discovery and Optimization, Preclinical Assessment (ADME/Toxicity), and Clinical Development.

2.1. Target Identification and Validation:

Identifying and validating the correct molecular target is the foundational step in drug discovery, and it is here that AI provides crucial early-stage acceleration [4].

Genomic and Proteomic Analysis: ML models (e.g., Random Forests, Support Vector Machines) are trained on large-scale datasets from genomics, transcriptomics, and proteomics to identify genes or proteins highly correlated with a specific disease state. DL models, such as Graph Neural Networks (GNNs), can model complex biological networks (protein-protein interactions, signaling pathways) to pinpoint novel, druggable nodes within a disease mechanism [5].

Predicting Target-Disease Associations: AI can cross-reference disease signatures with known drug actions and pathways, enabling the prioritization of thousands of potential targets into a manageable number of high-confidence candidates for experimental validation, significantly improving the 'hit rate' of the entire process [6].

2.2. Lead Discovery and Optimization:

Once a target is validated, the goal shifts to finding a small molecule (hit) that interacts with it and optimizing that molecule's properties (lead) into a clinical candidate.

Virtual Screening (VS): AI models enhance traditional VS by predicting the binding affinity and efficacy of millions of compounds against a target protein without the need for physical synthesis and testing. This is achieved using ML techniques trained on chemical descriptors and biological activity data, prioritizing compounds with the highest likelihood of becoming a 'hit' [7].

De Novo Drug Design: Generative AI models, including Variational Auto encoders (VAEs) and Generative Adversarial Networks (GANs), are used to design entirely novel drug-like molecules.

These systems learn the rules of medicinal chemistry from existing compound libraries and can generate new chemical structures tailored to specific desired properties (e.g., high affinity, low toxicity) [8].

Quantitative Structure-Activity Relationship (QSAR) Modeling: AI models establish a mathematical link between the chemical structure of a compound and its biological activity. This allows researchers to iteratively modify a lead compound to enhance potency and selectivity, a process known as lead optimization [9].

2.3. Drug Repurposing (Repositioning):

AI is a powerful tool for finding new uses for existing, approved drugs. Since these drugs have established safety profiles and pharmacokinetic data, repurposing them dramatically reduces development time and risk. AI algorithms analyze vast networks connecting drugs, targets, and diseases to predict previously unknown drug-disease relationships. For instance, AI has been used to identify drugs approved for one condition (e.g., arthritis) that may be effective against a completely different one (e.g., a rare cancer or an infectious disease) [10].

3. AI IN PRECLINICAL AND CLINICAL PHASES

3.1. Preclinical Assessment and ADME/Toxicity Prediction:

A primary cause of drug failure is poor ADME (Absorption, Distribution, Metabolism, and Excretion) properties and unforeseen toxicity. AI models are now integral to predicting these properties in silico before costly lab experiments are initiated [11].

Pharmacokinetic (PK) Prediction: ML models can accurately predict parameters like oral bioavailability, plasma protein binding, and half-life by analyzing the chemical structure of a molecule. This enables chemists to rapidly modify a lead compound to improve its drug-like characteristics [12].

Toxicity and Safety Prediction: AI systems, trained on public toxicity databases and proprietary preclinical data, are employed to predict various toxicological endpoints, including hepatotoxicity, cardiotoxicity, and mutagenicity. DL techniques, such as Convolutional Neural Networks (CNNs), are particularly effective at classifying molecular structures based on their predicted toxic risk [13]. This enhanced accuracy in early toxicity screening allows for the culling of high-risk compounds, saving significant time and animal resources.

3.2. AI in Clinical Trial Design and Management

The integration of AI in clinical trials aims to improve efficiency, reduce costs, and accelerate patient access to new treatments.

Optimized Patient Recruitment: Natural Language Processing (NLP) can analyze Electronic Health Records (EHRs) and clinical notes to automatically screen and match eligible patients to trials based on complex inclusion and exclusion criteria, significantly shortening the recruitment phase, which is often a major delay [14].

Trial Optimization: AI is used to model and predict the success probability of a clinical trial based on historical data, biomarker information, and trial design parameters. Furthermore, AI tools can optimize dosing schedules and select optimal trial sites, making the trials more efficient and less burdensome for participants [15].

Real-Time Monitoring and Data Analysis: Wearable devices and sensors generate continuous real-time data from patients. ML models analyze this massive data stream to detect subtle adverse events or treatment responses earlier than traditional periodic clinical visits, allowing for faster, data-driven decisions on trial modification or termination.

3.3. Personalized Medicine and Manufacturing

The impact of AI extends beyond the core discovery phases into patient-centric care and pharmaceutical production.

Personalized Dosing and Therapy: By integrating a patient's genetic data (pharmacogenomics), clinical history, and lifestyle factors, AI models can predict individual patient responses to a drug. This facilitates the shift from a "one-size-fits-all" approach to precision medicine, optimizing treatment efficacy and minimizing adverse effects for individual patients [16].

Smart Manufacturing (Pharma 4.0): AI is essential for continuous manufacturing processes. ML models can monitor sensor data from automated production lines in real-time to predict equipment failures, ensure continuous quality control, and optimize chemical reaction conditions, leading to greater efficiency and lower production costs [17].

4. CHALLENGES AND FUTURE OUTLOOK

While the promise of AI in the pharmaceutical industry is immense, several challenges must be addressed for its full potential to be realized [18].

4.1. Data Quality and Quantity Issues

Effective AI relies entirely on high-quality, vast, and well-annotated datasets. The data in pharmaceuticals is often sparse, heterogeneous, and siloed across different organizations.

Standardization: Lack of standardized data formats and protocols across institutions hinders data fusion and the creation of large, clean training sets.

Bias and Representativeness: Models trained on biased or unrepresentative historical patient data (e.g., lacking diversity across demographics) can perpetuate health disparities when applied to the broader population.

4.2. Model Interpretability (The Black Box Problem)

In a regulated industry, the ability to trace a prediction back to its source data and logic is paramount.

Lack of Transparency: Many powerful DL models, particularly those used for de novo design, function as "black boxes." This inherent difficulty in deciphering the model's internal workings makes it challenging for regulatory bodies to assess the scientific validity and risk profile of an AI-derived compound [19].

Reproducibility: Ensuring the reproducibility of predictions from models that are continuously learning or built on complex, proprietary architectures remains a significant technical and regulatory hurdle.

4.3. Regulatory and Ethical Hurdles:

The speed of AI innovation has outpaced traditional regulatory frameworks, creating uncertainty.

Evolving Guidance: Regulatory bodies (like the FDA and EMA) are actively developing guidance to address AI/ML applications, but definitive rules

for the approval of AI-generated molecules or AI-driven clinical decision systems are still evolving. This requires pharmaceutical companies to navigate an ambiguous landscape [20].

Intellectual Property and Liability: The use of generative AI in molecular design presents novel legal questions regarding who owns the intellectual property of a molecule designed by an algorithm. Furthermore, liability concerns arise if a decision informed by a flawed AI model leads to patient harm during a clinical trial.

Data Privacy: Utilizing patient-derived 'omics' data and real-world data from EHRs for training AI models necessitates strict adherence to data protection regulations (like GDPR and HIPAA) to maintain patient confidentiality and public trust.

CONCLUSION:

Artificial Intelligence represents a powerful, disruptive force poised to revolutionize the pharmaceutical industry. By augmenting human researchers with unprecedented analytical and predictive capabilities, AI is successfully tackling the long-standing problems of time, cost, and risk across the entire drug discovery and development spectrum. From identifying novel molecular targets via GNNs to optimizing clinical trial patient recruitment using NLP, AI is accelerating the pipeline and enabling the shift toward personalized medicine. Overcoming the challenges of data standardization, model interpretability, and regulatory clarity will solidify AI's role not just as a tool, but as an indispensable partner in the discovery of next-generation therapeutics.

ACKNOWLEDGMENTS:

The authors would like to acknowledge the pioneering work of all researchers who have contributed to the field of Artificial Intelligence and its application in drug discovery. The authors would also like to express their sincere gratitude to all the researchers and scholars whose work has contributed significantly to the development of this review. We also extend our heartfelt appreciation to Mr. Bapurao Patil, President, Madhavrao Patil College of Pharmacy, Murum, Mr. Sharan Basavaraj Patil, Governing body member of Madhavrao Patil Mahavidyalaya and staff for their continuous support and encouragement, which facilitated the successful completion of this work. Additionally, we thank colleagues and reviewers for their valuable feedback and suggestions that have helped improve the quality of this manuscript.

REFERENCE

1. DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics, 2016; 47: 20-33.
2. Waring, M. J., et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. Nature Reviews Drug Discovery, 2015; 14(7): 475-486.
3. Vamathevan, J., et al. Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 22(8), 2023.
4. Lyu, T., et al. Artificial intelligence in novel therapeutic target identification and validation. Trends in Pharmacological Sciences, 2024; 45(4): 277-290.
5. Zeng, W., et al. Graph Neural Networks in drug discovery. Drug Discovery Today, 2024; 29(1): 103734.
6. Xiong, Z., et al. Prediction of drug-target interactions based on graph convolutional

networks. Bioinformatics, 2020; 36(5): 1461-1468.
7. Gawehn, E., Hiss, J. A., & Schneider, G. Deep learning in drug discovery. Molecular Informatics, 2018; 37(12): 1700160.
8. Satz, A. L., et al. De novo molecular design using deep generative models. ACS Central Science, 2019; 5(5): 752-761.
9. Cherkasov, A., et al. The application of quantitative structure–activity relationships (QSAR) to modern drug discovery and environmental chemistry. Chemical Reviews, 2014; 114(5): 73-82.
10. T. Zaki, A. G., et al. Drug repurposing against COVID-19 using artificial intelligence: a review. International Journal of Pharmaceutical Sciences, 2023; 15(3): 67-80.
11. Ragoza, M., et al. Learning the ADMET landscape with deep reinforcement learning. ACS Central Science, 2020; 6(3): 353-363.
12. Xing, L., et al. Prediction of small molecule properties using graph attention networks. Journal of Chemical Information and Modeling, 2021; 61(1): 32-45.
13. Menden, M. P., et al. Community efforts to improve the prediction of clinical trial outcomes. Nature Communications, 2019; 10(1): 1802.
14. Corwin, D., et al. Using natural language processing to enhance clinical trial recruitment. JAMIA Open, 2021; 4(3): ooab033.
15. Rajkomar, A., et al. Scalable and accurate deep learning with electronic health records. Nature Medicine, 2018; 24(7): 1083-1089.
16. Ge, T., et al. Deep learning-based precision medicine for cancer. Artificial Intelligence in Medicine, 2023; 141: 102558.
17. Lee, S. L., et al. Continuous pharmaceutical manufacturing: smart processing using artificial intelligence. Journal of Pharmaceutical Sciences, 2022; 111(1): 66-78.
18. Paul, S. M., et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nature Reviews Drug Discovery, 2010; 9(3): 203-214.
19. Linardatos, P., Papastefanopoulos, E., & Kotsiantis, S. Interpretable machine learning: A perspective from medical and biological domains. Applied Sciences, 2020; 10(4): 1957.
20. Flaherty, M. J. The ethics of AI in drug discovery. Nature Medicine, 2023; 29(1): 41-43.

HOW TO CITE: C. Sadak Vali, Abdullah Khan, S. Siva Prasad, Mare Pratibha Bharathi, Reddy Nazemoon, Survi Swati Goud, Applications of AI in Pharmaceuticals and Drug Discovery: A Comprehensive Review, Int. J. of Pharm. Sci., 2025, Vol 3, Issue 11, 2880-2885. https://doi.org/10.5281/zenodo.17649085