



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



Review Article

AI in Drug Discovery and Development

Abhishek Sahu*, Prem Samundre, Dr. Jitendra Banweer

SAGE University Bhopal

ARTICLE INFO

Published: 15 May 2025

Keywords:

Artificial Intelligence, Drug discovery, Machine Learning, Deep Learning, Computational Drug Design, Virtual Screening, Drug Development Pipeline, Clinical Trial optimization

DOI:

10.5281/zenodo.15426727

ABSTRACT

Artificial intelligence (AI) is revolutionizing the landscape of drug discovery and development by accelerating timelines, reducing costs, and improving the precision of therapeutic design. From target identification to lead optimization and clinical trial design, AI-driven approaches—such as machine learning, deep learning, and natural language processing—are enhancing the efficiency and predictive power of each stage in the pharmaceutical pipeline. This review explores the current applications of AI across the drug development lifecycle, with a focus on virtual screening, de novo drug design, biomarker discovery, and patient stratification. We also discuss the challenges and limitations of integrating AI into biomedical research, including data quality, model interpretability, and regulatory considerations. By highlighting recent breakthroughs and emerging trends, this paper underscores the transformative potential of AI to redefine how new drugs are discovered, tested, and brought to market.

INTRODUCTION

BRIEF OVERVIEW OF DRUG DISCOVERY AND DEVELOPMENT TIMELINE

Drug discovery and development is a long and complex process that typically takes 10 to 15 years. It begins with identifying and optimizing potential drug compounds during the discovery phase, followed by preclinical testing in lab and animal models to assess safety and biological activity. If successful, the drug enters clinical trials in humans, progressing through three phases to

evaluate safety, efficacy, and optimal dosage. After completing trials, data is submitted for regulatory review. If approved, the drug enters the market, where ongoing post-marketing surveillance monitors long-term safety and effectiveness. (DiMasi et al., 2016).

CHALLENGES: HIGH ATTRIBUTION RATES, TIME CONSUMPTION, COMPLEX BIOLOGICAL DATA

Drug discovery and development face several major challenges that contribute to high attrition

***Corresponding Author:** Abhishek Sahu

Address: SAGE University Bhopal

Email ✉: abhisheksahu228489@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.



rates, extended timelines, and escalating costs. A significant number of drug candidates fail during clinical trials due to issues with safety, efficacy, or unforeseen side effects, leading to high failure rates. The entire process is time-consuming, often taking over a decade from initial discovery to market approval. Additionally, the complexity of biological systems and the vast, heterogeneous data generated from genomics, proteomics, and patient responses make it difficult to accurately predict drug behavior and outcomes (Scannell et al., 2012).

OBJECTIVE OF THE PAPER: TO REVIEW HOW AI IS REVOLUTIONIZING DRUG DISCOVERY AND DEVELOPMENT

The objective of this paper is to review how artificial intelligence (AI) is revolutionizing the field of drug discovery and development. It aims to explore the integration of AI-driven technologies across various stages of the drug development pipeline—from target identification and compound screening to clinical trial design and post-marketing surveillance. By highlighting current applications, recent advancements, and ongoing challenges, the paper seeks to provide a comprehensive understanding of AI's transformative impact on accelerating drug discovery, improving accuracy, and reducing costs.

ROLE OF AI IN DIFFERENT STAGES OF DRUG DISCOVERY

TARGET IDENTIFICATION AND VALIDATION

In drug discovery, target identification involves finding biological molecules that drive disease, while validation ensures these targets can be modulated for therapeutic benefit. AI plays a crucial role in both steps by analyzing large-scale biological data, such as genomics and proteomics,

to identify potential drug targets. Machine learning algorithms help predict the function of targets, and AI-based network analysis maps out biological pathways, pinpointing key proteins involved in disease. Additionally, AI models simulate drug-target interactions and predict potential side effects, enhancing target validation. By integrating multi-omics data and clinical insights, AI accelerates the discovery of druggable targets and improves the accuracy of the validation process (Rosenblatt et al., 2020).

DRUG SCREENING AND LEAD IDENTIFICATION

In the drug discovery process, drug screening involves testing large libraries of compounds to identify those that interact with a specific biological target, while lead identification focuses on finding promising compounds that can be optimized for further development. AI accelerates these stages by using machine learning models to predict how different compounds will interact with targets. Virtual screening, powered by AI, allows researchers to rapidly screen vast chemical libraries without the need for physical testing. Deep learning models are used to predict binding affinities and optimize the identification of potential drug leads. Additionally, generative models, such as variational autoencoders, can design novel molecules with desired properties. These AI-driven approaches significantly reduce the time and cost of drug screening, enabling the discovery of more effective leads (Vamathevan et al., 2019).

PRECLINICAL AND CLINICAL DEVELOPMENT

Preclinical and clinical development are the two major stages in the process of bringing a new drug to market. Preclinical development takes place before any testing in humans and involves



laboratory research using cell cultures (in vitro) and animal models (in vivo). These studies assess the drug's safety, toxicity, pharmacokinetics (how the drug moves through the body), and potential efficacy. The goal is to gather enough data to support an Investigational New Drug (IND) application, allowing the drug to proceed to human trials. Once the IND is approved, clinical development begins. This stage involves testing the drug in humans through several phases. Phase I focuses on evaluating safety and determining the appropriate dosage in a small group of healthy volunteers or patients. Phase II examines the drug's effectiveness and monitors for side effects in a larger group of patients. Phase III involves even more participants and compares the new drug to existing treatments, further confirming its safety and effectiveness.

AI TECHNIQUES AND TOOLS

Artificial Intelligence (AI) has become a transformative force in drug discovery and development, employing a variety of techniques such as machine learning (ML), deep learning (DL), natural language processing (NLP), and reinforcement learning (RL). Machine learning algorithms like Random Forests, Support Vector Machines, and Gradient Boosting are used for predicting drug-target interactions, toxicity, and pharmacokinetics. Deep learning models, including Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs), are especially effective in processing complex molecular data and structure-based predictions. NLP is widely used to extract biomedical insights from literature and clinical data, while reinforcement learning supports de novo drug design by generating novel molecules optimized for specific biological properties.

DATASETS AND PLATFORMS

COMMON PUBLIC DATABASES

In AI-based drug discovery, several public datasets and platforms are commonly used to train and validate models. ChEMBL and PubChem are major sources of bioactivity and chemical data, providing millions of compounds and assay results. DrugBank offers detailed drug-target interaction data, while the Protein Data Bank (PDB) contains 3D structures of proteins essential for structure-based drug design. BindingDB provides binding affinities between drugs and targets, and ZINC15 is widely used for virtual screening of purchasable compounds. Platforms like DeepChem, Open Targets, and RDKit support the integration and analysis of these datasets for AI applications in drug discovery.

PROPRIETARY VS OPEN SOURCE TOOLS (DeepChem, Atomwise, BenevolentAI)

In AI-driven drug discovery, tools are divided into open-source and proprietary categories, each with unique strengths. DeepChem is an open-source library that supports a wide range of machine learning and deep learning models for molecular property prediction, docking, and chemistry-related tasks. It's freely available, community-supported, and widely used in academic research. In contrast, Atomwise and BenevolentAI are proprietary platforms. Atomwise uses its patented AtomNet technology for structure-based virtual screening, targeting protein-ligand interactions at scale. BenevolentAI combines machine learning with biomedical knowledge graphs to accelerate drug discovery, from target identification to clinical development.

CLOUD COMPUTING AND BIG DATA PLATFORMS

Cloud computing and big data platforms play a crucial role in AI-powered drug discovery by enabling large-scale data storage, high-



performance computing, and collaborative research. Cloud platforms such as Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure offer scalable infrastructure for running machine learning models, virtual screening, and molecular simulations. These services also support specialized tools like AWS SageMaker and Google AI Platform for model training and deployment. On the big data side, platforms like Apache Hadoop, Apache Spark, and Databricks facilitate efficient processing of massive biomedical datasets, including genomics, chemical libraries, and real-world evidence. These technologies help researchers accelerate drug discovery workflows, reduce costs, and manage complex data pipelines across global teams.

CASE STUDIES

REAL-WORLD APPLICATION

AI is increasingly applied in real-world drug discovery with transformative results. AlphaFold by DeepMind has set a new standard in protein structure prediction, solving structures with near-laboratory accuracy and accelerating structure-based drug design (Jumper et al., 2021). In another case, Insilico Medicine successfully designed a novel preclinical drug candidate for pulmonary fibrosis using AI, reducing the time from target identification to preclinical testing to less than 18 months. These advancements show how AI can drastically shorten timelines and improve precision in drug development (Zhavoronkov, 2019).

AI SUCCESS STORIES DURING COVID-19

During the COVID-19 pandemic, AI played a critical role in accelerating vaccine and drug discovery. DeepMind's AlphaFold predicted the structure of SARS-CoV-2 proteins early in the pandemic, aiding researchers worldwide in understanding viral mechanisms (Jumper et al.,

2021). BenevolentAI used its AI platform to identify Baricitinib, a rheumatoid arthritis drug, as a potential COVID-19 treatment by analyzing biomedical data to uncover its anti-inflammatory and antiviral properties. It was later authorized for emergency use. AI also supported Moderna in rapidly designing its mRNA vaccine by analyzing viral genome sequences and optimizing mRNA constructs, contributing to the record-breaking vaccine development timeline. These successes demonstrate AI's power in accelerating pandemic responses.

REGULATORY AND ETHICAL CONSIDERATIONS

TRANSPARENCY & EXPLAINABILITY OF AI MODELS

Transparency and explainability are critical for building trust in AI models used in drug discovery and healthcare. Many advanced AI models, particularly deep learning networks, are often seen as "black boxes" due to their complex inner workings. This lack of interpretability can limit their adoption in regulated fields like pharmaceuticals, where understanding why a model makes a prediction is as important as the prediction itself. Techniques such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and attention mechanisms help provide insights into model decisions by highlighting which features most influenced the output. Efforts are also being made to develop inherently interpretable models and ensure compliance with regulatory expectations for transparency and reproducibility in AI-driven research.

REGULATORY PATHWAYS

The regulatory pathway for AI-driven drug discovery and development is still evolving, as regulatory bodies work to integrate advanced



technologies within existing frameworks. AI tools used in drug development must align with traditional approval processes like the Investigational New Drug (IND) and New Drug Application (NDA) pathways under the U.S. FDA, or equivalent processes through the European Medicines Agency (EMA). However, when AI is used for tasks such as target identification, compound screening, or biomarker discovery, questions around model transparency, data quality, and algorithm validation become critical. Regulatory agencies are beginning to address these issues through new guidance documents and reflection papers focused on AI and machine learning. The importance of robust governance, ethical AI design, and explainability is emphasized in regulatory science literature as essential to gain regulatory approval.

DATA PRIVACY, BIAS AND REPRODUCIBILITY

In AI-driven drug discovery, data privacy, bias, and reproducibility are critical ethical and technical concerns. The use of patient data—especially from electronic health records and genomics—raises privacy issues governed by regulations like HIPAA in the U.S. and GDPR in Europe. Bias can emerge from unbalanced training datasets, leading to inaccurate or unfair predictions, particularly for underrepresented populations. This can affect everything from drug response predictions to clinical trial design. Moreover, the reproducibility of AI models remains a challenge due to differences in data sources, model architectures, and lack of transparent reporting. Addressing these issues requires the adoption of standardized data protocols, model interpretability tools, and ethical AI guidelines to ensure reliable and equitable outcomes in drug development.

CONCLUSION & FUTURE DIRECTIONS

SUMMARY OF KEY BENEFITS AI BRINGS

AI brings several key benefits to drug discovery, including speed, cost-efficiency, and precision. It accelerates processes like target identification, compound screening, and drug repurposing, significantly reducing the time and cost traditionally required. AI also enables the analysis of massive and complex datasets—such as genomics, proteomics, and clinical records—to uncover hidden patterns and insights. Through predictive modeling, it improves decision-making in early-stage research and clinical trials.

HUMAN- AI COLLABORATION IN DRUG DESIGN

Human-AI collaboration in drug design leverages the strengths of both human expertise and artificial intelligence to enhance innovation and efficiency. AI assists researchers by rapidly generating and evaluating novel compounds, predicting drug-target interactions, and analyzing massive biological datasets. However, it is the human scientists who provide context, interpret results, ensure ethical integrity, and make critical decisions throughout the discovery process. This partnership allows for smarter, faster, and more targeted drug development. As emphasized by Eric Topol in *Deep Medicine*, the future of medicine is not about AI replacing humans, but about “deepening the human connection” by letting machines handle the data burden, so scientists and clinicians can focus on creativity, empathy, and strategic thinking.

OUTLOOK ON AI'S TRANSFORMATIVE ROLE IN PHARMA

AI's transformative role in the pharmaceutical industry is expected to reshape how drugs are discovered, developed, and delivered. AI can process vast datasets to accelerate drug discovery, optimize clinical trial designs, predict drug



efficacy, and personalize medicine, making the entire drug development process more efficient and cost-effective. In the future, AI will continue to integrate with technologies like precision medicine, genomics, and biometrics, enabling better-targeted therapies and more adaptive healthcare systems. AI's evolution will not only accelerate drug development timelines but also help create more individualized and effective treatments, addressing unmet medical needs and enhancing healthcare outcomes (Vamathevan et al., 2019).

REFERENCES

1. DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47, 20–33.
2. Scannell, J. W., Blanckley, A., Boldon, H., & Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery*, 11(3), 191–200.
3. Rosenblatt, M., Gopalakrishnan, V., & Luo, Z. (2020). Artificial intelligence in drug discovery: What's real, what's not, and where is it going? *Nature Biotechnology*, 38(6), 725–732.
4. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., & Ochoa, D. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, 18(6), 463–477.
5. Zhavoronkov, A. (2019). *Artificial Intelligence in Drug Discovery*. Academic Press. ISBN: 9780128161763.
6. Jumper, J. et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589.
7. Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books. ISBN: 9781541644632.
8. O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group. ISBN: 9780553418811.
9. Cincilla, G., Masoni, S., & Blobel, J. (2021). Individual and collective human intelligence in drug design: evaluating the search strategy. *Journal of Cheminformatics*, 13, 80. <https://doi.org/10.1186/s13321-021-00556-6>.
10. Vamathevan, J., Clark, J., Czodrowski, P., Dunham, I., & Ochoa, D. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, 18(6), 463–477.

HOW TO CITE: Abhishek Sahu*, Prem Samundre, Dr. Jitendra Banweer, AI in Drug Discovery and Development, *Int. J. of Pharm. Sci.*, 2025, Vol 3, Issue 5, 2510-2515. <https://doi.org/10.5281/zenodo.15426727>

