



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



## Review Paper

# A Comprehensive Review on Artificial Intelligence in Drug Discovery and Development

**Pranay Sawant\*, Pooja Surve, Dr Sunayana Ghodgaonkar, Suyash Salve, Afsana Shaikh, Taskin Shaikh**

*Shivajirao S. Jondhle college of pharmacy, Thane, India*

### ARTICLE INFO

Published: 28 Mar 2026

**Keywords:**

Artificial Intelligence (AI), drug discovery pipeline, Deep Neural Networks (DNNs), AlphaFold, target identification, de novo design

**DOI:**

10.5281/zenodo.19281998

### ABSTRACT

With clinical trial failure rates exceeding 90%, drug discovery and development is a complicated, expensive, and time-consuming process that usually takes 12 to 15 years and substantial financial investment to bring a single drug to market [1]. Inadequate target identification, poor therapeutic efficacy, and unexpected toxicity during late-stage trials are major reasons for attrition in traditional drug development [1-3]. These difficulties have forced the adoption of sophisticated computational and data-driven methods and revealed significant inefficiencies in traditional drug discovery pipelines [2, 3]. Pharmaceutical research can be accelerated, prediction accuracy can be increased, and experimental workload can be decreased thanks to artificial intelligence (AI), which combines machine learning (ML) and deep learning (DL) techniques [3-5]. Target identification and validation, hit identification via virtual screening, lead optimization, and prediction of absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties are examples of AI applications in drug discovery [6-9]. De novo molecular design and effective chemical space exploration have been made possible by sophisticated AI models like deep neural networks, convolutional neural networks, recurrent neural networks, generative adversarial networks, and variational autoencoders [6,10-12]. Additionally, structure-based drug design and early toxicity assessment have been greatly improved by AI-driven tools like graph-based learning models and protein structure prediction systems [10-13]. Beyond the early stages of discovery, AI has shown a significant impact on clinical trials and drug development by improving pharmacovigilance and regulatory compliance, optimizing trial design, and improving patient recruitment through electronic health record analysis [14-17]. Despite these developments, the broad use of AI in pharmaceutical research is still constrained by issues with data quality, model interpretability, ethical issues, regulatory

**\*Corresponding Author:** Pranay Sawant

**Address:** Shivajirao S. Jondhle college of pharmacy, Asangaon, Taluka- Shahapur District-Thane, State- Maharashtra, pin code-421601

**Email** ✉: [Pranaysawant.an@gmail.com](mailto:Pranaysawant.an@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.



acceptance, and high computational requirements [15–17]. All things considered, artificial intelligence is a paradigm shift in drug discovery and development, with the potential to lower costs, expedite development, and promote the development of safer, more efficient, and customized therapeutic interventions [1–17].

## INTRODUCTION

Most people agree that one of the most difficult and resource-intensive procedures in the pharmaceutical industry is drug research and development. To get a single therapeutic agent from initial discovery to market approval, the traditional drug development process usually takes 12 to 15 years and costs between 1.5 and 2.6 billion US dollars [1]. With over 90% of drug concepts failing during clinical development, especially in the later stages of clinical trials, the likelihood of success is still astonishingly low despite these large investments [1,2]. Inadequate target selection, poor therapeutic efficacy, unattractive pharmacokinetic features, and unexpected toxicity are the main causes of these high attrition rates, all of which point to basic inefficiencies in conventional drug development methods [2,3]. Drug development has traditionally been mostly dependent on empirical, costly, and time-consuming trial-and-error methods. While hit discovery and lead optimization rely on intensive experimental screening and chemical synthesis efforts, target identification and validation are frequently predicated on scant biological knowledge [3]. Furthermore, the large and intricate biological datasets produced by contemporary high-throughput technologies are difficult for traditional experimental techniques to properly examine. Because of this, promising drug candidates often fail at later stages of development, resulting in large financial losses and delayed patient access to viable medications [4].

A paradigm shift toward data-driven drug discovery has been fueled in recent years by the quick expansion of processing power, the accessibility of vast amounts of biological data, and advancements in algorithmic development. Many of the drawbacks of conventional pharmaceutical research can be addressed by artificial intelligence (AI), which includes machine learning (ML) and deep learning (DL) approaches [5]. Predictive modeling and well-informed decision-making throughout the drug discovery and development process are made possible by AI systems' ability to recognize intricate patterns in massive, high-dimensional datasets [6]. AI models can quickly produce insights that would otherwise require substantial experimental validation by learning from past chemical, biological, and clinical data [7]. All things considered, artificial intelligence offers the ability to shorten development times, cut expenses, and increase the success rate of therapeutic innovation, making it a revolutionary approach to drug research and development. The future of pharmaceutical research is anticipated to be further shaped by AI technologies as they continue to develop and interact with cutting-edge disciplines including genomics, systems biology, and precision medicine [3,17].

## 2. Fundamentals of Artificial Intelligence in Drug Discovery

Computational systems created to mimic human cognitive processes including learning, thinking, and decision-making are referred to as artificial intelligence. Algorithms may learn patterns from past data without explicit programming thanks to machine learning, a fundamental component of artificial intelligence. ML approaches are widely used in pharmacological research to predict pharmacokinetic characteristics, toxicity, and molecular activity [3]. To carry out classification and regression tasks, such as figuring out if a



compound is active or inactive against a biological target, supervised learning models use labelled datasets. Compound clustering and illness stratification are made easier by unsupervised learning techniques, which examine unlabeled information to find hidden structures and trends. Through reward-based feedback mechanisms, reinforcement learning enhances decision-making. It is especially useful for molecular scaffold optimization and de novo molecular design [4].

Based on deep neural networks with several hidden layers, deep learning is a sophisticated type of machine learning that can directly extract intricate, non-linear properties from unprocessed data. Activity prediction, toxicity modeling, and property estimate are common applications of deep neural networks [5]. Convolutional neural networks are specialized architectures that are commonly used for binding affinity prediction in two-dimensional and three-dimensional molecule representations. They are intended to capture

spatial and structural information. Recurrent neural networks are frequently used to model protein–ligand interactions and generate chemical structures represented as SMILES strings because they are designed for sequential data. Furthermore, generative deep learning models have become increasingly important in contemporary drug development. While variational autoencoders support lead optimization and chemical space exploration by encoding molecular information into latent spaces and decoding it into new chemical structures, generative adversarial networks create novel chemical entities through a competitive learning process involving a generator and a discriminator [6]. In QSAR modeling, virtual screening, and ADMET prediction, conventional machine learning algorithms like support vector machines and random forests are still crucial [7].

**Table 1. Fundamentals of Artificial Intelligence**

Concept	Description	Application in Drug Discovery
<b>Machine Learning (ML)</b>	Learns from data without explicit programming	Activity prediction, QSAR
<b>Supervised Learning</b>	Uses labeled data	Active vs inactive molecule prediction
<b>Unsupervised Learning</b>	Finds hidden patterns in unlabeled data	Clustering similar compounds
<b>Reinforcement Learning (RL)</b>	Learns via rewards and penalties	De novo molecular design
<b>Deep Learning (DL)</b>	Uses deep neural networks	Feature extraction from raw data
<b>CNNs</b>	Capture spatial features	Binding affinity prediction
<b>RNNs</b>	Handle sequential data	SMILES generation, protein–ligand modeling
<b>GANs</b>	Generator vs discriminator model	Novel drug-like molecule generation
<b>VAEs</b>	Latent space representation	Lead optimization
<b>SVM &amp; Random Forest</b>	Traditional ML models	ADMET prediction, virtual screening

### 3. Data Requirements for AI-Driven Drug Discovery

The availability of sizable, varied, and high-quality datasets is essential to the efficacy of AI



models in drug discovery. The big data platform that underpins AI-driven pharmaceutical research is defined by volume, velocity, and variety [5]. Predictive modeling is based on chemical and molecular data, such as SMILES representations, molecular fingerprints, and physical characteristics like solubility and lipophilicity. Target identification and validation depend on omics data, which includes genomes, proteomics, transcriptomics, and metabolomics. These data offer profound insights into disease mechanisms [8]. High-throughput screening procedures produce biological assay data that provide important insights into biological activity and compound–target interactions. Molecular docking research and structure-based drug design depend heavily on structural data, such as three-dimensional protein structures and protein–ligand complexes. Clinical trial optimization and drug repurposing tactics are further supported by clinical data from electronic health records, clinical trials, and real-world evidence [9]. AI model performance is still constrained by issues such data heterogeneity, a lack of high-quality labeled datasets, privacy issues, and restricted data sharing because of proprietary restrictions.

**Table 2. Types Of Data Sets**

Method	Description
<b>Structure-Based Virtual Screening (SBVS)</b>	Uses 3D protein structures and AI-enhanced docking
<b>Ligand-Based Virtual Screening (LBVS)</b>	QSAR and DL models identify similar active compounds

**Table 1. Types Of Datasets**

Category	Examples
<b>Chemical Data</b>	PubChem, ChEMBL, ZINC
<b>Genomic Data</b>	GEO, TCGA
<b>Structural Data</b>	Protein Data Bank (PDB)
<b>Drug Data</b>	DrugBank, TTD

#### 4. Computational Tools, Databases, and Platforms

AI-based drug discovery has advanced quickly because to strong computational tools and publicly available information. Accurate prediction of molecular attributes is made possible by graph neural networks, which offer an effective framework for describing molecules as graphs with atoms as nodes and bonds as edges [10]. A deep learning-based technology called DeepTox was created to predict the toxicity profiles of thousands of compounds, making it easier to rule out dangerous medication candidates early on [11]. By precisely identifying three-dimensional protein structures from amino acid sequences, Google DeepMind's AlphaFold has transformed protein structure prediction and greatly advanced structure-based drug design and target validation [12]. While genomic and expression datasets are accessible through platforms like GEO and TCGA, public databases like PubChem, ChEMBL, and ZINC offer a wealth of chemical and bioactivity data. DrugBank and the Therapeutic Target Database provide extensive drug and target data, while the Protein Data Bank is the main source of structural information. AI-driven pharmaceutical research is supported by software frameworks like DeepChem, ORGANIC, and IBM Watson that combine machine learning, reinforcement learning, generative modeling, and natural language processing [13].

#### 5. Applications of AI in Drug Discovery



Target identification and validation, a stage where poor choices often lead to downstream failures, is where artificial intelligence plays a crucial role. AI models find potential druggable targets and disease-associated biomarkers that might be missed by traditional methods by analysing high-dimensional multi-omics datasets [8]. By building and analysing intricate biological interaction networks, network-based techniques make it easier to identify important genes and proteins that are crucial for the advancement of disease. Target validation is strengthened and structure-based drug design is supported by accurate protein structure prediction using AlphaFold [12]. AI significantly speeds up the discovery of compounds that can interact with biological targets in hit identification and virtual screening. Three-dimensional target structures are used in structure-based virtual screening, and AI-enhanced molecular docking increases the precision of binding affinity estimates. While deep learning-based QSAR models allow for effective screening of extremely vast chemical libraries, ligand-based virtual screening uses known active compounds to find new molecules with similar features [14]. In conventional drug development, lead optimization and ADMET prediction are significant bottlenecks. In order to reduce late-stage failures, AI models analyse chemical structures to forecast absorption, distribution, metabolism, excretion, and toxicity qualities early on. While some models forecast drug-drug interactions and adverse events, tools like DeepTox produce thorough toxicity profiles [11]. Additionally, artificial intelligence speeds up molecular dynamics simulations, making it possible to accurately forecast drug stability and behaviour under physiological settings in much shorter computational times [15]. By forecasting a compound's effectiveness against several targets, artificial intelligence (AI) further enhances polypharmacology, a technique that is especially useful for treating complicated

disorders involving multiple biological pathways. Furthermore, ideal reaction paths and bond disconnections are predicted using AI-driven retrosynthesis and chemical synthesis planning, which expedites laboratory procedures and lessens experimental effort [16].

## 6. AI in Drug Development and Clinical Trials

The majority of the time and expense involved in bringing a drug to market is spent on drug development and clinical trials, where artificial intelligence has shown a considerable impact. In order to optimize trial design, including endpoint selection, dosing schedules, and inclusion or exclusion criteria, AI-driven models examine genomic data, electronic health records, and past trial results [17]. AI-based site selection and feasibility analysis increase trial success rates and recruitment effectiveness. Additionally, natural language processing methods make it possible to quickly identify qualified patients from unstructured clinical data, greatly speeding up the recruitment process. Additionally, safety monitoring, efficacy prediction, and regulatory compliance are all improved by artificial intelligence. AI models assist in forecasting adverse events and treatment outcomes by combining preclinical data, past trial findings, and empirical knowledge. AI facilitates quicker and more precise safety evaluations in pharmacovigilance by automating the processing of adverse drug reaction data and safety paperwork. Clinical trial dataset's dependability and integrity are further guaranteed by automated data validation and error detection [15].

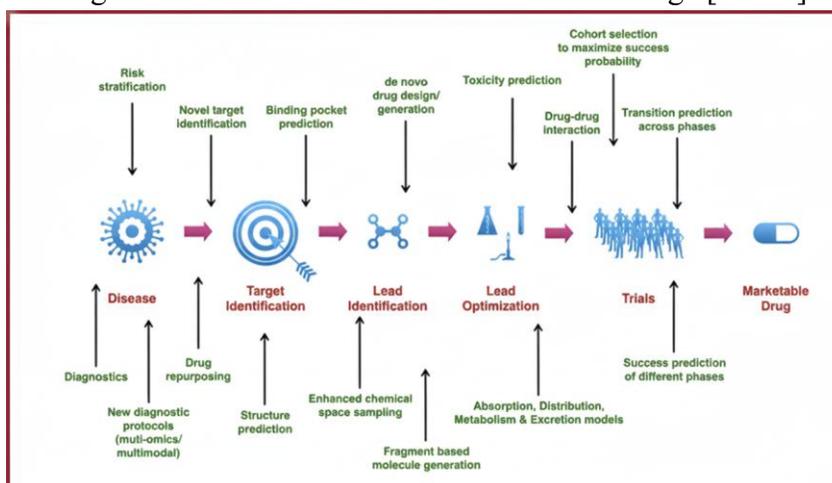
AI is also essential for site selection and feasibility assessment, both of which are important factors in determining trial success. AI algorithms can determine high-performing trial sites and forecast recruiting feasibility by examining patient demographics, investigator performance, geographic illness incidence, and past enrollment



data. This data-driven strategy addresses one of the most frequent reasons for clinical trial failure by increasing enrollment rates, reducing delays, and improving overall trial efficiency [15].

In conclusion, by improving trial design, speeding up patient recruitment, strengthening safety monitoring, and cutting development costs and delays, artificial intelligence has become a

revolutionary force in drug development and clinical trials. It is anticipated that the incorporation of AI techniques into clinical research will become more crucial in enhancing trial success rates and promoting the creation of safe, efficient, and customized treatments as these approaches continue to advance and regulatory frameworks change [14–17].



**Fig 1. AI In Drug Development**

## 7. Challenges and Limitations

Artificial intelligence has the potential to revolutionize drug research, but there are a number of obstacles to overcome. High-quality labeled datasets are hard to come by, data sources are inconsistent, and data sharing is limited because of privacy and proprietary issues. Both the high computational cost of training complicated models and the black-box aspect of deep learning models, which restricts interpretability and confidence among regulators and physicians, present computational problems. Data privacy, bias, and fairness are still major ethical issues, especially when dealing with sensitive genetic and medical data. Furthermore, a significant obstacle to wider adoption is the lack of experts with training in the nexus of AI and biological sciences [15,16].

The interpretability, computational requirements, and regulatory acceptability of AI-driven methods represent yet another significant barrier. Many sophisticated AI models, especially deep learning

systems, function as "black boxes," making it challenging to evaluate or explain their predictions. This undermines confidence among researchers, medical professionals, and regulatory bodies. Smaller businesses are also hampered by the expensive infrastructure and computing costs involved in developing and maintaining sophisticated AI models. Clinical translation is made more difficult by the absence of established regulatory frameworks for assessing AI-based predictions as well as ethical issues with data security and privacy. To properly utilize AI in pharmaceutical research, these constraints must be addressed by enhanced data governance, the creation of explainable AI models, interdisciplinary cooperation, and regulatory harmonization [15–17].

## 8. Future Perspectives

It is anticipated that intimate human-AI collaboration will be a part of drug research in the future, where computational forecasts will supplement human expertise. It is expected that developments in explainable AI would increase regulatory acceptability and openness. Pharmaceutical innovation is expected to be further accelerated by the integration of AI with

cutting-edge technologies including generative models, large-scale data analytics, and sophisticated computational platforms. Furthermore, the development of precision and personalized medicine approaches will be made easier by AI-driven integration of multi-omics and clinical data [14,17].

**Table 4. Future Of AI**

Area	Future Outlook
<b>Human-AI collaboration</b>	AI will assist researchers in faster and better decision-making
<b>Explainable AI</b>	Transparent models will improve trust and regulatory acceptance
<b>Generative AI</b>	Faster design of novel drug molecules
<b>Multi-omics integration</b>	Better target identification and disease understanding
<b>Personalized medicine</b>	Patient-specific treatment strategies
<b>Clinical trials</b>	Improved trial design and patient recruitment

The combination of emerging technologies and interdisciplinary data sources is also anticipated to expedite future advancements in AI-driven drug discovery. massive-scale biological data analytics, generative modeling, and sophisticated computational platforms combined with AI will make it possible to efficiently explore massive chemical and biological domains. Furthermore, it is anticipated that the integration of multi-omics data with empirical clinical evidence would propel the advancement of precision and personalized medicine strategies, enabling the customization of treatments to specific patient profiles. AI has the potential to significantly influence the development of the next generation of safer, more efficient, and patient-specific medicinal approaches as data exchange programs, computing infrastructure, and interdisciplinary training continue to grow [8,14–17].

## CONCLUSION

The integration of artificial intelligence into drug discovery and development has fundamentally reshaped the traditional pharmaceutical research paradigm, which has long been characterized by

high costs, lengthy timelines, and extremely high failure rates. Conventional drug development processes rely heavily on empirical experimentation and trial-and-error approaches, often resulting in the late-stage failure of promising candidates due to poor efficacy, toxicity, or unfavorable pharmacokinetic profiles. AI, through machine learning and deep learning techniques, offers a data-driven alternative capable of analysing vast and complex chemical, biological, and clinical datasets to generate accurate predictions at every stage of the drug development pipeline.

In the early phases of discovery, AI enhances target identification and validation by integrating multi-omics data and biological interaction networks to uncover novel disease-associated biomarkers and druggable targets. During hit identification and lead optimization, AI-assisted virtual screening, QSAR modeling, and generative deep learning methods accelerate the exploration of chemical space and enable the design of novel drug-like molecules with optimized properties. AI-based ADMET and toxicity prediction further reduce attrition rates by identifying unsafe or



ineffective compounds at an early stage, thereby conserving both time and financial resources.

Addressing these challenges will require the development of explainable AI models, standardized validation procedures, improved data governance policies, and closer collaboration between data scientists, biologists, clinicians, and regulatory authorities. As these barriers are gradually overcome, artificial intelligence is expected to become a cornerstone of future pharmaceutical innovation, enabling faster, more economical, and more personalized drug development. Ultimately, the widespread implementation of AI has the potential to deliver safer and more effective therapies, reduce global healthcare burdens, and usher in a new era of precision medicine.

## REFERENCES

1. Abbas MKG, Rassam A, Karamshahi F, Abunora R, Abouseada M. The Role of AI in Drug Discovery. *ChemBioChem*. 2024;25:e202300816.
2. Winkler DA. Use of Artificial Intelligence and Machine Learning for Discovery of Drugs for Neglected Tropical Diseases. *Front Chem*. 2021;9:614073.
3. Vora LK, Gholap AD, Jetha K, Thakur RRS, Solanki HK, Chavda VP. Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design. *Pharmaceutics*. 2023;15(7):1916.
4. Anonymous. The Role of Artificial Intelligence in Modern Drug Discovery and Development. [Journal Unknown]; PDF ID: d46347rfdy.
5. Zhu H. Big Data and Artificial Intelligence Modeling for Drug Discovery. *Annu Rev Pharmacol Toxicol*. 2020;60:573–589.
6. Koçak M, Akçalı Z. The published role of artificial intelligence in drug discovery and development: a bibliometric and social network analysis from 1990 to 2023. *J Cheminform*. 2025;17:71.
7. Singh S, Kumar R, Payra S, Singh SK. Artificial Intelligence and Machine Learning in Pharmacological Research: Bridging the Gap Between Data and Drug Discovery. *Cureus*. 2023;15(8):e44359.
8. Jiménez-Luna J, Grisoni F, Weskamp N, Schneider G. Artificial intelligence in drug discovery: recent advances and future perspectives. *Expert Opin Drug Discov*. 2021;16(9):949–959.
9. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers*. 2021. doi:10.1007/s11030-021-10217-3.
10. Bali A, Bali N. Role of artificial intelligence in fast-track drug discovery and vaccine development for COVID-19. In: *Novel AI and Data Science Advancements for Sustainability in the Era of COVID-19*. Elsevier; 2022. p.201–230.
11. Liebman M. The role of artificial intelligence in drug discovery and development. *Chem Int*. 2022;Jan–Mar:16–17. doi:10.1515/ci-2022-0105.
12. Ocana A, Pandiella A, Privat C, Bravo I, Luengo-Oroz M, Amir E, Gyorffy B. Integrating artificial intelligence in drug discovery and early drug development: A transformative approach. *Biomark Res*. 2025;13:45.
13. Niazi SK, Mariam Z. Artificial intelligence in drug development: reshaping the therapeutic landscape. *Ther Adv Drug Saf*. 2025;16:1–24. doi:10.1177/20420986251321704.
14. Fu C, Chen Q. The future of pharmaceuticals: Artificial intelligence in drug discovery and development. *J Pharm Anal*. 2025;15:101248. doi:10.1016/j.jpha.2025.101248.



15. Blanco-González A, Cabezón A, Seco-González A, Conde-Torres D, Antelo-Riveiro P, Piñeiro Á, Garcia-Fandino R. The role of AI in drug discovery: Challenges, opportunities, and strategies. *Pharmaceuticals*. 2023;16:891.
16. Ozaybi MQB, Madkhali ANM, Alhazmi MAM, Faqihi HMA, Alanazi MM, Siraj WHY, et al. The role of artificial intelligence in drug discovery and development. *Egypt J Chem*. 2024;67(SI):1541–1547. doi:10.21608/ejchem.2024.337877.10835.
17. Patel V, Shah M. Artificial intelligence and machine learning in drug discovery and development. *Intelligent Medicine*. 2022;2:134–140. doi:10.1016/j.imed.2021.10.001.

**HOW TO CITE:** Pranay Sawant, Pooja Surve, Dr Sunayana Ghodgaonkar, Suyash Salve, Afsana Shaikh, Taskin Shaikh, A Comprehensive Review on Artificial Intelligence in Drug Discovery and Development, *Int. J. of Pharm. Sci.*, 2026, Vol 4, Issue 3, 3778-3786, <https://doi.org/10.5281/zenodo.19281998>

