International Journal of Research in Engineering, Science and Management
Volume 8, Issue 9, September 2025
https://www.ijresm.com | ISSN (Online): 2581-5792

AI Technology in Drug Discovery

J. Rekha Rani[1], Roshan Kumar[2*], Muskan Shaikh[3], Prerana Prayasi[4], Zaid Naik[5], Mousumi Mondal[6]

***Abstract*: Artificial intelligence (AI) has the potential to revolutionize the drug discovery process, offering improved efficiency, accuracy, and speed. However, the successful application of AI is dependent on the availability of high-quality data, the addressing of ethical concerns, and the recognition of the limitations of AI-based approaches. In this article, the benefits, challenges, and drawbacks of AI in this field are reviewed, and possible strategies and approaches for overcoming the present obstacles are proposed. The use of data augmentation, explainable AI, and the integration of AI with traditional experimental methods, as well as the potential advantages of AI in pharmaceutical research, are also discussed. Overall, this review highlights the potential of AI in drug discovery and provides insights into the challenges and opportunities for realizing its potential in this field.**

***Keywords*: Drug discovery, Data augmentation, AI in pharmaceutical research.**

1. Introduction to AI

A. What is Artificial Intelligence

- Artificial Intelligence or AI is a branch of computer science focused on building systems that can perform tasks typically requiring human intelligence.
- These tasks include learning from experience, reasoning through logic, understanding language, and even making decisions.
- AI systems use algorithms and models especially machine learning and deep learning to identify patterns, adapt to new data, and improve performance over time.
- In simple terms, AI allows computers to 'think' and 'learn' in ways that mimic human behaviour. It is already impacting many industries from finance to transportation and today, we'll explore its powerful role in transforming drug discovery and development [9].

B. AI in Healthcare

- Natural Language Processing is used to analyse clinical notes, patient history, and electronic health records (EHRs).
- Robotic surgeries assisted by AI offer precision, minimal invasiveness, and faster recovery.
- AI supports drug discovery, genomics, personalized medicine, and public health surveillance.
- AI is playing crucial role in the healthcare industry by improving diagnosis, treatment planning, and patient care.
- Used in medical imaging such as (CT, MRI, X-ray) to detect diseases like cancer, stroke, and fractures with high accuracy.
- Predictive analytics helps identify at-risk patients and prevent complications. AI-powered chatbots and virtual assistants improve patient engagement and reduce administrative burden [5].

C. AI in Pharma

- The integration of AI enables faster drug target identification, molecule generation, ADMET prediction, and clinical trial optimization.
- *Key AI technologies used include*:
  - Machine Learning (ML) [13]
  - Deep Learning (DL) [10], [11]
  - Natural Language Processing (NLP)
  - Reinforcement Learning (RL)
- In the pharmaceutical industry, AI is transforming how drugs are discovered, designed, and developed.
- Traditional drug discovery is expensive, time-consuming, and has a high failure rate, AI helps overcome these limitations.
- AI can analyse vast datasets, identify hidden patterns, and predict biological behaviour of compounds more accurately than conventional methods.

D. The Promise of Artificial Intelligence

Artificial intelligence represents a paradigm shift in drug discovery, potentially reducing timelines by 50–70% while increasing success rates through more accurate prediction of clinical outcomes.

- AI is not simply a new tool, but a transformative approach that fundamentally changes how we discover and develop drug

E. Drug Discovery

- *Target Identification*: Process of identifying and validating a biological target relevant to the disease of interest.
- *Hit Discovery*: Screening compounds to find those that interact with the target (hits).
- *Lead Optimization:* Refining promising compounds to improve potency, selectivity, and pharmacological properties.
- *Preclinical Development*: Testing compounds in

*Corresponding author: shubhamraaz071@gmail.com

laboratory and animal models to assess safety and efficacy.

- *Clinical Phases*: Progressive human trials to determine safety, efficacy, dosage, and monitor adverse effects [6], [14].

F. Key Limitations in Drug Discovery

Time & Cost: Lengthy development timelines (10-15 years) with extremely high costs ($2-3 billion per approved drug).

High Attrition Rate: Over 90% of drug candidates fail during development, with most failures occurring in late-stage clinical trials.

Data Complexity: Inability to efficiently process vast biological datasets and extract meaningful patterns.

Limited Prediction Accuracy: Poor translation between in vitro, animal, and human studies, leading to unexpected clinical failures.

Inefficient Screening: Traditional screening approaches test limited compound libraries and miss potential therapeutic molecules.

G. Popular AI Tools & Platforms

AlphaFold (DeepMind): Revolutionary AI system for highly accurate protein structure prediction, transforming structural biology research [2].

Insilico Medicine's GENTRL: Generative Tensorial Reinforcement Learning system for de novo drug design that produces novel molecules [3].

Atomwise AtomNet: Deep convolutional neural network for virtual screening and structure-based drug design.

Benevolent AI: Knowledge graph and machine learning platform for target identification and drug repurposing [4].

H. Target Identification: From Years to Days

Global: Insilico Medicine (Hong Kong) used generative adversarial networks (GANs) to identify a novel target for idiopathic pulmonary fibrosis in 46 days – a process traditionally taking 3-5 years. The target was validated in vivo within 9 months (Nature Biotechnology, 2019).

India: Tata Medical Center (Kolkata) collaborates with IISc Bangalore on Project AIRA (AI for Rare Diseases), using NLP to mine Indian clinical records and genomic databases. In 2023, it identified 3 novel targets for oral cancer prevalent in betel nut chewers – a population underrepresented in Western datasets [20].

2. AI – Clinical Trial Design

A. AI-Enabled Patient Cohort Identification, Site Selection, and Trial Optimization

1) Patient Cohort Identification

AI analyses EHRs and omics data to identify optimal patient populations and enhance trial stratification.

AI analyses EHRs, genomic data, and biomarkers to identify suitable participants.

Improves inclusion/exclusion criteria and ensures diverse representation.

Site Selection Optimization: Algorithms predict high-performing sites based on historical data, demographics, and infrastructure.

AI ranks clinical trial sites based on historical performance, patient availability, and compliance data.

2) AI-Enhanced Clinical Trials

30-40% reduction in patient recruitment time.

15-25% decrease in overall trial costs.

- *Trial Protocol Optimization*: Machine learning models simulate outcomes to design optimal trial protocols. Helps in selecting endpoints, dosage, and control groups effectively.
- *Adaptive Trial Design*: AI enables real-time adjustments to trial parameters based on interim data. Reduces cost and improves trial efficiency and success rates.
- *Predictive Analytics*: Forecasts patient responses, dropout likelihood, and adverse events. Supports risk-based monitoring and proactive decision-making.

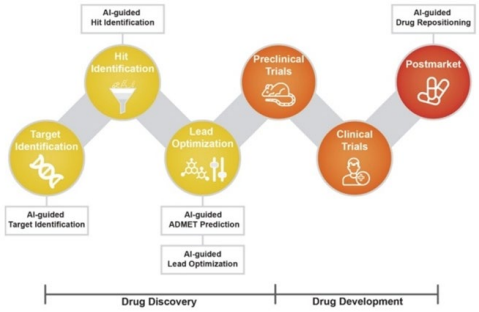

Fig. 1. AI in clinical trials and drug development

3. AI Technology in Drug Discovery

A. How AI Improves Chemical Properties, Potency, and Safety Profiles

- *Multi-Objective Optimization*: AI simultaneously balances potency, selectivity, safety, and pharmacokinetic properties using sophisticated algorithms.
- *Generative Models*: Deep learning generates novel molecular structures with desired properties, expanding chemical space exploration.
- *Predictive QSAR Models*: AI predicts structure-activity relationships to guide optimization with minimal experimental validation.
- *Synthetic Accessibility*: AI evaluates synthesizability of proposed compounds, prioritizing candidates that are chemically feasible.
- *Accelerated Optimization*: Reinforcement learning algorithms efficiently navigate chemical space to identify optimal compounds with fewer iterations.

B. Sources of Pharmaceutical Data

- *Omics Databases*: Genomic, proteomic, and metabolomic data repositories like GenBank, UniProt, and Metabo Lights.
- *Literature and Patents*: PubMed, Google Scholar,

patent databases – containing valuable research findings and IP information.

C. How AI Identifies and Validates Novel Drug Targets Using Omics and Big Data

- *Multi-omics Data Integration*: AI algorithms analyse genomics, proteomics, and metabolomics data to identify potential therapeutic targets.
- *Network Biology Analysis*: ML models map complex biological interactions to discover disease-associated pathways and targets.
- *In Silico Target Validation*: AI predicts target druggability, efficacy, and potential off-target effects before lab experiments.
  - *literature to identify Literature Mining & Knowledge Graphs*: NLP extracts insights from scientific novel targets and validate hypotheses.
  - *Biomarker Identification*: ML algorithms identify biomarkers that predict drug response for targeted therapies.
  - *Generative Models*: AI systems that can create entirely new molecular structures without relying on existing compound libraries.
  - *Deep Reinforcement Learning*: Rewards models for generating molecules with specific properties like binding affinity, synthetic accessibility, and drug-likeness.
  - *Variational Autoencoders (VAEs)*: Encode molecules into a continuous latent space, enabling exploration of chemical space and novel compound generation.
  - *Generative Adversarial Networks (GANs)*: Use competing neural networks to create highly realistic molecular structures meeting multiple optimization criteria.
  - *Multi-Objective Optimization*: Simultaneously optimizing diverse molecular properties (potency, selectivity, solubility, synthesizability).

D. The Key AI Application in Drug Development

1) *Target Identification & Validation*: Predicting disease related genes and proteins.
2) *De Novo Drug Design*: Generating novel compounds using generative models [17].
3) *Virtual Screening*: Identifying promising drug candidates from large compound libraries.
4) *ADMET Prediction*: Assessing Absorption, Distribution, Metabolism, Excretion and Toxicity early in silico.
5) *Clinical Trial Optimization*: Patient recruitment, stratification, and response prediction [15].
6) *Drug Repurposing*: Identifying new uses for existing drugs using AI pattern recognition [7], [8].

E. Preclinical Development

- AI predicts in vivo toxicity, bioavailability, and pharmacokinetic behavior.
- Simulates animal studies to reduce reliance on animal testing.
- Speeds up formulation development by optimizing excipients and delivery systems.
- *Intelligent Animal Model Selection*: AI analyses disease mechanisms to identify most relevant and predictive animal models, increasing translational success.
- *Automated Histopathology*: Machine learning algorithms analyse tissue samples and microscopy data with greater speed and consistency than manual methods.
- *Advanced Toxicity Prediction*: Deep learning models predict organ-specific toxicities and adverse effects before animal testing begins [21], [25].
- *In Silico Trials*: Virtual simulations of preclinical studies to optimize experimental design and reduce animal usage.
- *Digital Biomarker Development*: AI identifies novel measurable indicators of drug effects that can translate from animal models to humans.

4. AI Technology in Post Market Surveillance

A. AI for Pharmacovigilance, Adverse Event Detection, and Drug Repurposing

- *Advanced Pharmacovigilance*: AI systems analyse vast amounts of real-world data to detect safety signals and adverse events that might be missed by traditional methods.
- *Diverse Data Sources*: Integration of electronic health records, insurance claims, social media, wearable devices, and patient-reported outcomes for comprehensive monitoring.
- *Signal Detection*: Machine learning algorithms identify potential safety concerns by detecting statistical anomalies and previously unknown drug-event associations.
- *Drug Repurposing*: Analysis of real-world usage patterns reveals off-label benefits, leading to new indications for existing drugs.
- *Regulatory Integration*: FDA and EMA are developing frameworks for using AI-derived real-world evidence in regulatory decision-making [16].

5. Notable Case Studies

A. Target Identification: From Years to Days

- *Global*: Insilico Medicine (Hong Kong) used generative adversarial networks (GANs) to identify a novel target for idiopathic pulmonary fibrosis in 46 days – a process traditionally taking 3–5 years. The target was validated in vivo within 9 months (Nature Biotechnology, 2019).
- *India*: Tata Medical Center (Kolkata) collaborates with IISc Bangalore on Project AIRA (AI for Rare Diseases), using NLP to mine Indian clinical records

and genomic databases. In 2023, it identified 3 novel targets for oral cancer prevalent in betel nut chewers – a population underrepresented in Western datasets.

B. *Virtual Screening at Scale*

Global: Atomwise (USA) screened 7 billion compounds in 48 hours using convolutional neural networks to find Ebola therapeutics, leading to two candidates with nanomolar binding affinity.

1) *Generative Chemistry: Digital Molecule Design*

Global: UK's Ex Scientia designed the first AI-generated drug candidate DSP-1181 (for OCD) in 12 months – 4x faster than industry standards. The compound entered Phase I trials in 2020.

Atom wise & Ebola: Used AI to identify two promising compounds in less than a day.

Insilico Medicine (2020): Designed a novel fibrosis drug candidate in just 46 days using generative AI.

Benevolent AI & COVID-19: Identified baricitinib as a potential treatment, which later entered clinical trials.

6. Future Prospects of AI in Drug Discovery and Development

A. *Next-Generation Drug Discovery*

- *AI + Quantum Computing*: Predicting molecular behaviour with quantum precision.
- *AI-designed novel molecules*: Generative AI tools will create new drug candidates never seen before.
- *Target identification*: Improved understanding of complex diseases through systems biology and AI modelling [1].

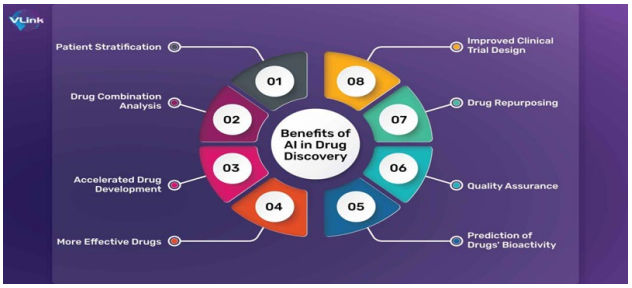

Fig. 2. Benefits of AI in Drug discovery

B. *AI in Precision Medicine*

- *Patient stratification*: AI-driven biomarker discovery to identify responders/non-responders.
- *Personalized therapy*: Custom drug regimens based on genomic and phenotypic data.

7. Keys Take aways

A. *Key Insights*

- *Accelerated Discovery:* AI reduces drug discovery timelines by 50–70% while maintaining or improving quality outcomes [22].
- *Improved Success Rates:* Machine learning models enhance prediction accuracy for drug candidates, reducing clinical trial failures [23].
- *Novel Target Discovery:* AI enables identification of previously unexplored targets through complex data integration [24].

B. *Recommendations*

- *Invest in Interdisciplinary Teams:* Build teams with expertise in AI, chemistry, biology, and clinical research.
- *Prioritize Data Quality:* Establish robust data standards and curation practices to ensure AI model reliability.
- *Develop Strategic Partnerships:* Collaborate with AI technology providers and academic institutions to access cutting-edge capabilities.

8. Conclusion

- AI's future in pharma is not just complementary.
- It's transformational from ideation to market launch.
- AI promises to reduce timeframes, enhance precision, and increase success rates—paving the way for a smarter, patient-centric, and globally accessible healthcare ecosystem.
- AI-based methods are being adopted in the health care industry where low-cost, intelligent, and flexible methods are affecting areas such as drug design, support for clinical decision making, diagnosis, prevention, and making clinical recommendations.
- AI applications were previously thought to be inferior to experimental high-throughput screening, combinatorial chemistry, and other technical drivers. It was difficult to create new chemical entities using computer programs, with desired features from the ground up, potentially even better than a human expert.
- AI is now being utilized to create representations of trials that allow for data categorization and, ultimately, the development of predictive models.[18,19]
- AI has the potential to revolutionize the drug discovery process, offering improved efficiency and accuracy, accelerated drug development, and the capacity for the development of more effective and personalized treatments . However, the successful application of AI in drug discovery is dependent on the availability of high-quality data, the addressing of ethical concerns, and the recognition of the limitations of AI-based approaches.
- The utilization of AI in virtual screening and the meticulous design of drugs exemplifies its capacity to refine drug development methodologies.
- This computational efficiency extends to polypharmacology, chemical synthesis, and drug repurposing, showcasing AI′s transformative potential in improving healthcare outcomes on a global scale.
- The emergence of machine learning, a key subset of AI, facilitates the development of models capable of learning and making predictions autonomously, further enriching the drug discovery landscape.

- AI′s contribution to drug development is undeniably pivotal, offering a pathway to more efficient, effective, and personalized medical treatments. As the pharmaceutical industry continues to embrace AI, overcoming its inherent challenges will be crucial in realizing its full potential, ultimately leading to groundbreaking advancements in healthcare and therapeutic discovery.
- AI holds immense promise to bridge the gap between molecular complexity and therapeutic innovation. Its continued evolution will likely define the next era of personalized, inflammation-targeted cancer treatments.
- The integration of AI in drug design and discovery has ushered in a transformative era for the pharmaceutical industry. AI models contribute significantly to key processes including target identification, drug design, lead identification, drug repurposing, lead optimization, and toxicity prediction, substantially accelerating the development pipeline.
- AI-powered approaches are proving invaluable in identifying drug repurposing candidates, an especially vital area in the context of emerging diseases and pandemics. By enabling more accurate predictions, analysing large-scale data sets, and simulating complex molecular interactions, AI holds the potential to significantly reduce the time, cost, and failure rates traditionally associated with drug development. However, challenges remain, including data quality, interpretability of models, and the need for collaboration between AI experts and domain-specific researchers.
- Nevertheless, as AI continues to evolve and mature, its role in drug discovery is poised to enhance the precision, efficiency, and success of developing new therapeutic agents, ultimately improving patient outcomes globally.

References

[1] Mak, K.-K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. *Drug Discovery Today, 24*(3), 773–780.

[2] DeepMind (2021). AlphaFold: A solution to a 50-year-old grand challenge in biology. https://deepmind.com/research/highlighted-research/alphafold

[3] Insilico Medicine (2023). AI-powered end-to-end drug discovery platforms. https://insilico.com

[4] Benevolent AI. AI-driven drug pipeline and platform tools. https://www.benevolent.com

[5] U.S. FDA (2021). Artificial Intelligence and Machine Learning in Software as a Medical Device. https://www.fda.gov/media/145022/download

[6] Aliper, A., & Zhavoronkov, A. (Eds.) (2020). *Artificial Intelligence in Drug Discovery*. Springer Nature.

[7] Sarker, I. H. (2021). *Machine Learning in Drug Discovery: Methods and Applications*. Springer.

[8] Farghali, H.; Canová, N.K.; Arora, M. The Potential Applications of Artificial Intelligence in Drug Discovery and Development. *Physiol. Res.* 2021, *70* (Suppl. S4), S715–S722.

[9] Paul, D.; Sanap, G.; Shenoy, S.; Kalyane, D.; Kalia, K.; Tekade, R.K. Artificial intelligence in drug discovery and development. *Drug Discov. Today* 2021, *26*, 80–93.

[10] Zhuang, D.; Ibrahim, A.K. Deep learning for drug discovery: A study of identifying high efficacy drug compounds using a cascade transfer learning approach. *Appl. Sci.* 2021, *11*, 7772.

[11] Gawehn, E.; Hiss, J.A.; Schneider, G. Deep Learning in Drug Discovery. *Mol. Inform.* 2016, *35*, 3–14.

[12] Wess, G.; Urmann, M.; Sickenberger, B. Medicinal Chemistry: Challenges and Opportunities. *Angew. Chem. Int. Ed.* 2001, *40*, 3341–3350.

[13] Chen, R.; Liu, X.; Jin, S.; Lin, J.; Liu, J., Machine learning for drug-target interaction prediction. *Molecules* 2018, *23*, 2208.

[14] Sliwoski, G., Kothiwale, S. Meiler, J., Computational methods in drug discovery *Pharmacol. Rev.* 2014; 66:334-395.

[15] Woo, M. An ai boost for clinical trials *Nature.* 2019; 573, S100.

[16] Khalil Zadeh, N. · Sepehri, M.M. · Farvaresh, H. Intelligent sales prediction for pharmaceutical distribution companies: a data mining-based approach *Math. Probl. Eng.* 2014; 2014.

[17] Atance, S.R. · Diez, J. V. Engkvist, O., De novo drug design using reinforcement learning with graph-based deep generative models, 2021.

[18] Kempt, H. · Nagel, S.K. Responsibility, second opinions and peer-disagreement: ethical and epistemological challenges of using ai in clinical diagnostic contexts *J. Med. Ethics.* 2022; 48:222-229.

[19] Schneider, G. An insight into artificial intelligence in drug discovery: an interview with professor gisbert Schneider *Expert Opin. Drug Discov.* 2021; 16:933-935.

[20] Schneider, G. · Schneider, P. Macromolecular target prediction by self-organizing feature maps *Expert Opin. Drug Discov.* 2017; 12:271-277.

[21] Lin, A. · Giuliano, C.J. · Palladino, A., "Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials," *Sci. Transl. Med.* 2019; 11.

[22] Doni Dermawan & Nasser Alotaiq (2025), *"From Lab to Clinic: How Artificial Intelligence (AI) Is Reshaping Drug Discovery Timelines and Industry Outcomes".*

[23] Sydney Anuyah, Mallika K. Singh & Hope Nyavor (2024), *"Advancing Clinical Trial Outcomes Using Deep Learning and Predictive Modelling: Bridging Precision Medicine and Patient-Centered Care".*

[24] Aaron Wenteler, Claudia P. Cabrera, Wei Wei, Victor Neduva & Michael R. Barnes (2024) authored *"AI approaches for the discovery and validation of drug targets".*

[25] Avila A. M., Bebenek I., Bonzo J. A., Bourcier T., Davis Bruno K. L., Carlson D. B., et al. (2020). An FDA/CDER perspective on nonclinical testing strategies: classical toxicology approaches and new approach methodologies (NAMs). *Regul. Toxicol. Pharmacol*. 114, 104662.