International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Revolutionizing Drug Discovery: The Role of Artificial Intelligence

Bhavinkumar Shah

Abstract: ***This article explores the integration of artificial intelligence in the drug discovery process. It highlights the stages of drug development, the role of AI in enhancing efficiency, and the tools used in AI-assisted drug discovery. The article also discusses the challenges and prospects in this field.***

Keywords: Artificial intelligence, drug discovery, drug development, Machine learning

1. Introduction

The development of a new drug is a very lengthy and expensive process. The average time for a new drug to enter the market is about 12-15 years. The expected cost to develop a new drug-including capital costs and expenditure on drugs that fail to reach the market-has been estimated to range from less than $1 billion to more than $2 billion[1]. The new drug discovery consists of five main stages: 1) the Early drug discovery and development stage, in which basic research is performed to understand the mechanism leading to disease and propose a possible target {eg protein}; the drug discovery stage, during which the scientist looks for molecules {small molecules and biologicals} or other therapeutic strategies that interfere or cure the investigated disease or at least alleviate the symptoms; 2) the preclinical development stage, before testing the drug in people, the researcher must find whether it has the potential to cause severe harm and also toxicity, and this is done via invitro and in-vivo studies; 3) the clinical stage this stage investigate the drug candidate in human to make sure they are safe and effective; 4) the FDA review stage: FDA review teams thoroughly examine all of the submitted data related to the drug or device and make a decision to approve or not to approve it; 5) FDA post-market safety monitoring, FDA monitors all drug and device safety once products are available for use by the public[2].

Artificial intelligence is a field that combines computer science and robust datasets to enable problem-solving. It also encompasses machine learning and deep learning sub-fields, frequently mentioned with artificial intelligence[3]. Artificial intelligence-integrated drug discovery and development has accelerated the growth of the pharmaceutical sector, leading to a revolutionary change in the pharmaceutical industry. The article provides valuable insights into how AI can revolutionize drug discovery, significantly reducing time and cost, and highlights its growing importance in the pharmaceutical industry.

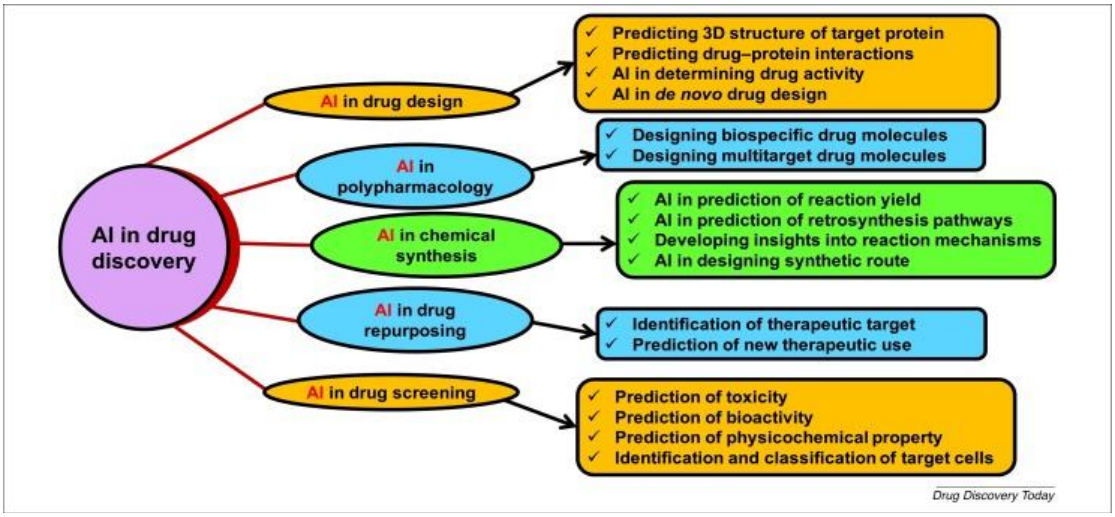
AI: networks and tools

The foundation of AI is based on machine learning [ML]. ML uses algorithms that recognize patterns within a data set and learn how to make predictions and recommendations by processing data and experience. A subset of machine learning, also known as deep learning, uses artificial neural networks (ANN) to mimic the human brain's learning process. These comprise a set of interconnected sophisticated computing elements involving 'perceptons' analogous to human biological neurons, mimicking the transmission of electrical impulses in the human brain[4]. There are different types of artificial neural networks used in machine learning:

- Feed-forward neural networks: here, the information only moves in one direction: forward from the model's input layer to its output layer, without ever traveling backward to be reanalyzed by the model.
- Multi-layer perceptron network: Multi-layer perceptron (MLP) supplements a feed-forward neural network. It consists of three types of layers—the input layer, output layer, and hidden layer. The MLP has applications including pattern recognition, optimization aids, and process identification[5].
- Convolutional neural network (CNN): it is a type of feed-forward neural network. They are used in image and video processing, biological systems, and complex brain functions[6].
- Recurrent neural network (RNN): RNNs are artificial neural networks whose connections include loops, meaning the model both moves data forward and loops it backward to run again through previous layers. RNNs help predict a sentiment or an ending of a sequence, like a large sample of text, speech, or images. They can do this because each input is fed into the model by itself and in combination with the preceding input[7].

AI in drug discovery

Finding a successful new drug is the most challenging part of drug development, mainly due to chemical space. Chemical space is a set of every possible chemical compound, including every known drug molecule and those that have yet to be discovered; the total number of compounds that make up the chemical space is estimated to be in the order of 10^{60} molecules.



The chemical space can give us information about molecules within the space, helping us look for bioactive compounds, and thus, virtual screening helps select molecules for further testing. AI, with the help of ZINC (An open resource for virtual screening of compounds), PubChem (the most extensive collection of freely accessible chemical and bio-activity information), and Drug Bank (A knowledge-base of drugs, drug action, and drug targets), provide a better profile analysis, faster elimination of non-lead compounds and selection of drug molecule, with reduced expenditure[8].

A quantitative structure-activity relationship (QSAR) based model can quickly predict many compound or simple physio-chemical parameters, such as log P or log D. QSAR modeling tools have been utilized for the identification of potential drug candidates and have evolved, such as linear discriminant analysis (LDA), support vector machine (SVM), random forest (RF) and decision trees, applied to speed up QSAR analysis[9].

AI in drug screening

Drug screening is the process by which potential drugs are identified and optimized before selecting a candidate drug to progress to clinical trials. It can involve screening large libraries of chemicals for a particular biological activity in high-throughput screening assays[10]. Different AI-based tools can be used to determine the physicochemical properties of the drug, like solubility, partition coefficient (log p), degree of ionization, and intrinsic permeability[11]. Neural network based on the ADMET predictor and ALGOPS program are used to predict the solubility and lipophilicity of various compounds[12]. Six Artificial intelligence methods, namely, Support vector machine, k-nearest neighbor, Probabilistic neural network, Artificial neural network, Partial least square, and Linear discriminant analysis, were used to predict compound absorption of compounds[13].

List of AI-based software for drug discovery, development, and analysis

Reference	Description
AlphaFold2	Deep learning-based model for 3D structure prediction of proteins from amino acid sequences
DeepChem	A deep learning library for drug discovery and computational chemistry
DeepBind	A computational tool to analyze binding between the protein and DNA/RNA
DeepBar	A method for accurate and fast prediction of binding free energy
Deep-Screening	Web server based in deep learning for virtual screening of compounds
DeepScreen	High-performance drug target interaction
DeepConv-DTI	A convolutional neural network-based model for predicting drug-target interactions
DeepPurpose	A Deep learning library for drug-target interaction, drug-drug interaction, protein-protein interaction, and protein function prediction
DeepTox	A deep learning model for toxicity prediction of chemical compounds

List of Compound and Bio-activity databases

Reference	Description
PubChem	Largest collection of freely accessible chemical and bio-activity information
ChEMBL	A large-scale bioactivity database for drug discovery
DrugBank	A knowledge-base of drugs, drug actions, and drug targets
ZINC	An open resource for virtual screening of compounds
BindingDB	A database of measuring binding affinity between target and the drug
ADME	An online database for pharmacokinetic information
STITCH	An integrated database of chemical-protein interactions
SIDER	Marketed medicines and their recorded adverse drug reactions
GDSC	Drug response data and genomic biomarkers
PDBBind	A comprehensive collection of binding affinities for the protein-ligand complexes in the Protein Data Bank (PDB)
canSar	Cancer translational research and drug discovery knowledgebase

Companies implementing AI in drug discovery and development[14].

1) Exscientia: A UK-based company has pioneered AI in small molecule drug design. The company helps develop novel therapeutic antibodies through generative AI design. It has partnered with big pharma giants like Bristol-MyersSquibb, Sanofi, GSK, and pathAI on drug discovery projects. In September 2023, they announced a new collaboration with Merck KGaA, Darmstadt, Germany, to discover novel small molecule drug candidates across oncology, neuroinflammation, and immunology[15].
2) Atomwise: it specializes in using AI in small-molecule drug discovery. AtomNet is a deep learning-driven computational platform for structure-based drug design developed by Atomwise. Its library includes more than 3 trillion synthesizable compounds. In 2022, it signed an agreement with Sanofi. Sanofi will pay Atomwise $20 million to identify, synthesize, and advance lead compounds for up to five targets exclusive to Sanofi[16].
3) Recursion Pharmaceuticals: A clinical-stage biotechnology company specializing in drug discovery through machine learning using its proprietary recursion operating system. It now has several compounds in phase 1 and 2 studies, including a small molecule therapeutics for cavernous cerebral malformation and another for neurofibromatosis type 2.
4) Iktos: They specialize in using AI technology for ligand and structure-based de novo drug design. The company's offerings include 1. dockAI, A new technology combining docking with a state-of-the-art active learning methodology.2. Makya: SaaS platform for AI-driven new drug design focused on multi-parametric optimization (MPO).3. Spaya: discover and prioritize the best synthetic route to compound using AI[17].
5) BenevolentAI: The clinical-stage drug discovery company developed an AI-enabled platform to accelerate drug discovery. BenAI engine is the core foundation of their technology that turns fragmented scientific data into actionable insight[18]. It is designed to ingest, organize, and leverage biomedical data for a nuanced understanding of disease biology[18]. The company has collaborated with multiple big pharmaceutical companies, including AstraZeneca, Merck, and Novartis. Its alliance with AstraZeneca centers around identifying novel drug targets for chronic kidney disease (CKD) and idiopathic pulmonary fibrosis (IPF). That collaboration has resulted in the selection of three novel targets for drug development. In September 2023, their collaboration with Merckutilizes Benevolent AI's end-to-end AI platform capabilities to deliver novel drug candidates, initially for three targets in oncology, neurology, and immunology.
6) CytoReason: it is a leading technology company developing computational disease models. The company collects proprietary data from pharmaceutical companies to stimulate human disease-tissue by tissue, cell by cell. Sanofi uses Cytoreason's AI platform in inflammatory bowel disease (IBD) to identify patient subtypes and pair them with IBD targets. Pfizer has used CytoReason's biological models in its research to enhance the understanding of the immune system as it develops innovative drugs for immune-mediated and immuno-oncology diseases.

Challenges in Adopting AI

2. Data

a) Availability: The topmost barrier is data availability. ML and deep learning models require large datasets to classify or predict different tasks accurately[19]. Most progressions are made where ML has large datasets to enable more complex and precise algorithms[20].
b) Privacy and security: most hospitals don't like to share data between them, as they are considered the property of each hospital. Healthcare data is sensitive and a frequent target for breaches. With the development of AI comes an additional concern regarding data privacy, as individuals may mistake artificial systems for humans and allow further unconscious data collection[21].
c) Data quality: Data quality measures how well a dataset meets criteria for accuracy, completeness, validity, consistency, uniqueness, timeliness, and fitness for purpose and it is critical to all data governance initiatives within an organization[22]. An issue with any of the above can hamper the decision-making process, leading to incorrect analysis.
d) Longevity: Patient data is estimated to have a half-life of roughly four months, implying that some predictive models may not be as successful at anticipating future outcomes as replicating those of the past[23].

3. Development Challenges:

3.1 Bias: Bias is prejudice or favoritism in thoughts or judgment. Bias in AI refers to the potential for unfairness or discrimination in the outcomes produced by AI systems. As AI technologies become more prevalent, IT professionals must understand and address bias. Failure to recognize and mitigate bias can have significant ethical, social, and legal implications.

3.2 Black-box: One widespread criticism of AI technologies is the 'black box' problem. Deep learning algorithms are often unable to provide detailed explanations for their predictions[24]. This becomes an issue because if the prediction is wrong, the system can't justify the reason.

3.3 Overfitting: Overfitting may occur where the system learns relationships between patient variables and outcomes that are not relevant. It is the result of having too many variable parameters relative to outcomes. Thus, the algorithm predicts using inappropriate features. The algorithm may work within the training dataset but give inaccurate results when predicting future outcomes.

4. Implementation

4.1 The foremost barrier is the need for more empirical evidence proving the efficacy of AI-based interventionin prospective clinical trials. Empirical research remains more extensive and primarily pertains to AI in the general workforce, not its effect on patient outcomes[25]. Randomized controlled trials are considered the gold standard in medicine

but must be improvedto prove the efficacy of AI in healthcare[26].

4.2 The second barrier would be the successful implementation of AI in physician workflow. For AI-based interventions, be that to do with scan interpretation or navigating electronic medical records, they must speed up rather than slow down physicians. This includes time and resources spent training physicians and healthcare providers to use the technology. To this date, instances of successful integration of AI into clinical care still need to be improved and remain primarily in trial stages[27].

Milestones in AI-Enabled Drug Discovery

Far from being a distant sci-fi future, AI-enabled drug discovery is already here. A non-exhaustive list of historic milestones in the field includes the following achievements:

- In early 2020, Exscientia announced the first-ever AI-designed drug molecule to enter human clinical trials.
- In July 2021, an AI system by DeepMind called AlphaFold predicted the protein structures for 330, 000 proteins, including all 20, 000 proteins in the human genome. The AlphaFold Protein Structure Database has since expanded to include over 200 million proteins, covering nearly all cataloged proteins known to science.
- In February 2022, Insilico Medicine reported the start of Phase I clinical trials for the first-ever AI-discovered molecule based on an AI-discovered novel target—all done at a fraction of the time and cost of traditional preclinical programs.
- In January 2023, AbSci became the first entity "to create and validate *de novo* antibodies *in silico*" using generative AI.
- In February 2023, the FDA granted its first Orphan Drug Designation to a drug discovered and designed using AI; Insilico Medicine plans to begin a global Phase II trial for the drug "early" this year.

According to Boston Consulting Group, as of March 2022, "biotech companies using an AI-first approach [had] more than 150 small-molecule drugs in discovery and more than 15 already in clinical trials

5. Conclusion

Integrating artificial intelligence (AI) in drug discovery has the potential to revolutionize the pharmaceutical industry and the healthcare sector. AI has untapped potential in healthcare. If successfully implemented, AI could relieve the workload for healthcare professionals and increase the quality of work produced by reducing error and increasing precision. It could grant patients more responsibility in their health management and reduce unnecessary hospital admissions. It could also extend the scope of medical knowledge, improving current clinical recommendations. It can accelerate drug development, improve clinical trial designs, predict drug bioactivity, and ensure quality assurance. Using the latest AI-based technologies will not only speed up the time needed for the product to come to the market. Still, it will also improve the quality of products, the overall safety of the production process, and the utilization of available resources, along with being cost-effective, thereby increasing the importance of automation. A significant worry regarding incorporating these technologies is the job losses that would follow, and the strict regulations needed to implement AI. However, these systems are intended only to make work more accessible rather than to replace humans altogether. The article concludes that AI holds tremendous potential in drug discovery, offering significant improvements in efficiency, cost-effectiveness, and quality. It emphasizes the need for further research and development to fully harness AIs capabilities in this field.

References

[1] https://www. cbo. gov/publication/57126
[2] https://www. fda. gov/patients/learn-about-drug-and-device-approvals/drug-development-process
[3] https://www. ibm.com/topics/artificial-intelligence
[4] Beneke F., Mackenrodt M. -O. Artificial intelligence and collusion. *IIC Int. Rev. Intellectual Property Competition Law*. 2019; 50: 109-134.
[5] Da Silva I. N. Springer; 2017. Artificial Neural Networks.
[6] Hänggi M., Moschytz G. S. Springer Science & Business Media; 2000. Cellular Neural Networks: Analysis, Design and Optimization.
[7] https://www. mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai
[8] Mak, Kit-Kay, and Mallikarjuna Rao Pichika. "Artificial intelligence in drug development: present status and future prospects. " *Drug Discovery Today* 24.3 (2019): 773-780.
[9] Zhang, Lu, et al. "From machine learning to deep learning: progress in machine intelligence for rational drug discovery. " *Drug discovery today* 22.11 (2017): 1680-1685.
[10] https://www. nature.com/subjects/drug-screening#: ~: text=Drug%20screening%20is%20the%20process, in%20high%2Dthroughput%20screening%20assays.
[11] Zang, Qingda, et al. "In silico prediction of physicochemical properties of environmental chemicals using molecular fingerprints and machine learning. " *Journal of chemical information and modeling* 57.1 (2017): 36-49.
[12] Lusci, Alessandro, Gianluca Pollastri, and Pierre Baldi. "Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules. " *Journal of chemical information and modeling* 53.7 (2013): 1563-1575.
[13] Kumar, Rajnish, et al. "Prediction of human intestinal absorption of compounds using artificial intelligence techniques. " *Current drug discovery technologies* 14.4 (2017): 244-254.
[14] https://www. drugdiscoverytrends.com/10-pioneering-companies-implementing-ai-in-drug-discovery-development-and-beyond/
[15] https://s28. q4cdn.com/460399462/files/doc_news/2023/08/Exscientia-Announces-AI-Drug-Discovery-Collaboration-with-Merck-KGaA-Darmstadt-Germany-2023. pdf
[16] https://www. atomwise.com/2022/08/17/atomwise-signs-strategic-multi-target-research-collaboration-with-sanofi-for-ai-powered-drug-discovery/
[17] https://iktos. ai/offerings/
[18] https://www. benevolent.com/benevolent-platform/

[19] . S. Ji, Q. Gu, H. Weng, Q. Liu, P. Zhou, Q. He, R. Beyah, T. Wang, De-health: all your online health information are belong to us. arXiv preprint.
[20] Lopez K, Fodeh S, Allam A et al. Reducing annotation burden through multimodal learning. Frontiers In Big Data 2020; 3: 8. 10.3389/fdata.2020.00019.
[21] Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. ArtifIntell Med
[22] https://www. ibm.com/topics/data-quality#: ~: text=the%20next%20step-, What%20is%20data%20quality%3F, governance%20initiatives%20within%20an%20organization.
[23] Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology.
[24] Wang F, Preininger A. AI in health: state of the art, challenges, and future directions.
[25] Sun TQ, Medaglia R. Mapping the challenges of artificial intelligence in the public sector: evidence from public healthcare.
[26] Ramesh, A. N., et al. "Artificial intelligence in medicine. " *Annals of the Royal College of Surgeons of England* 86.5 (2004): 334.
[27] Stewart, Jonathon, Peter Sprivulis, and Girish Dwivedi. "Artificial intelligence and machine learning in emergency medicine. " *Emergency Medicine Australasia* 30.6 (2018): 870-874.