Review Article

AI Use in Drug Discovery and its Challenges

Sakshi Wankhade, Mohit Kanherkar*, Purva Amzare, Sanika Ambekar, Dr. Harigopal Sawarkar
Dr. Rajendra Gode College of Pharmacy, Amravati, Andhra Pradesh, India.
***Corresponding author's E-mail:** mohitkanherkar@gmail.com

Received: 08-07-2026; **Revised:** 23-09-2026; **Accepted:** 29-09-2026; **Published online:** 20-10-2026.

ABSTRACT

Artificial intelligence (AI) is revolutionizing medication development by increasing the process's speed, effectiveness, and affordability. Natural language processing (NLP), machine learning (ML), and deep learning (DL) are three ways artificial intelligence (AI) may evaluate massive datasets to rapidly find pharmacological targets, forecast compound efficacy, and improve medication design. By forecasting pharmacokinetics, toxicity, and adverse effects, it also expedites lead identification. Additionally, it enhances clinical trial design by means of improved data analysis and patient selection. This review lists the main advantages of artificial intelligence (AI) in drug development, including increased productivity, enhanced precision, and lower risks. Nevertheless, there are still difficulties, such as poor data quality, opaque models, and regulatory worries.

Keywords: Artificial Intelligence; Drug Discovery; Machine Learning; Deep Learning; Virtual Screening; Generative AI; Drug Design.

INTRODUCTION

Data digitization has significantly increased in the pharmaceutical industry in recent years. But this change makes it harder to gather, evaluate, and use that data to solve complicated clinical issues [1]. The ability of AI to process vast amounts of data more efficiently and automatically is what propels its adoption [2]. AI is a technology-driven system that mimics human intellect using sophisticated tools and networks; it is not intended to take the place of people [3,4]. AI makes use of tools and software that can decipher input data, learn from it, and decide on their own how to proceed in order to accomplish particular objectives. According to this review, its uses in the pharmaceutical industry are growing quickly. The rapid advancements in AI-driven automation are predicted to drastically alter the work culture of society, according to the McKinsey Global Institute [5,6].

A key component of improving healthcare is the lengthy, intricate, and resource-intensive process of drug research and development. The discipline has evolved greatly from the first use of natural items and conventional treatments to the extremely advanced methods employed today. In the past, the majority of medicine research was empirical, using natural chemicals and trial and error. Scientists were able to find, produce, and test molecules with particular therapeutic promise thanks to a more methodical approach made possible by advancements in chemistry, biology, and pharmacology.

Target identification and validation, lead discovery, preclinical testing, and clinical trials are typically the main phases of the contemporary drug development pipeline. Before a possible medication may be administered to patients, each step is crucial to confirming its safety and efficacy. Even with advances in technology, the process is still time-consuming and frequently involves billions of dollars in research and investment. Challenges are further compounded by high failure rates, especially during clinical studies.

Target identification and validation, lead discovery, preclinical testing, and clinical trials are typically the main phases of the contemporary drug development pipeline. Before a possible medication may be administered to patients, each step is crucial to confirming its safety and efficacy. Even with advances in technology, the process is still time-consuming and frequently involves billions of dollars in research and investment. Challenges are further compounded by high failure rates, especially during clinical studies.

AI in drug discovery:

More than ten million molecules make up the huge chemical space, which facilitates the development of many potential drugs [7]. However, the lack of cutting-edge technologies slows down and increases the expense of drug research; AI can help with this [8]. AI has the ability to find hit and lead compounds, expedite the validation of therapeutic targets, and aid in drug design optimization [7,9]. Different Applications of AI in drug discovery are illustrated in figure 1.

The size, speed, diversity, and unpredictability of the data present significant data-related obstacles for AI, notwithstanding its benefits. Millions of chemicals are frequently included in drug research datasets used by pharmaceutical corporations, which may be too large for conventional machine learning methods to manage. Simple physicochemical parameters, like log P or log D, or the activity of a large number of compounds can be swiftly predicted using quantitative structure-activity relationship (QSAR) models. These models can't yet accurately forecast complicated biological characteristics, such a compound's effectiveness or its adverse effects. Furthermore, there are issues with quantitative structure-activity relationship-based models, such as limited training datasets,

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.

experimental data errors, and a dearth of experimental validation. These issues can be addressed by using big data analysis to assess the efficacy and safety of pharmacological compounds using emerging AI approaches like deep learning (DL) and associated modelling tools. A QSAR machine learning challenge was funded by Merck in 2012 to investigate the advantages of deep learning in drug discovery. For 15 distinct datasets pertaining to drug candidate absorption, distribution, metabolism, excretion, and toxicity (ADMET), DL models had much superior predictive performance than conventional machine learning techniques [10,11]. The virtual chemical space is vast and can be visualized like a geographical map, showing the distribution of molecules and their properties. Mapping this chemical space helps identify the positions of molecules, which is useful for discovering bioactive compounds. Virtual screening (VS) uses this information to select promising molecules for further testing. Several open-access chemical space databases are available, including PubChem, ChemBank, DrugBank, and ChemDB. Several in silico techniques are employed to virtually screen molecules from chemical spaces, such as ligand-based and structure-based methods. These techniques expedite the removal of non-lead compounds, enhance profiling, assist in the selection of possible therapeutic candidates, and lower overall expenses [7]. The physical, chemical, and toxicological characteristics of compounds are taken into consideration by drug design algorithms like LOMB matrices and molecular fingerprint identification in order to find appropriate lead candidates [12].

A molecule's intended chemical structure can be predicted using a variety of methods, including in silico techniques, compound generation methods, molecular similarity, and prediction models [9,13]. The DeepVS system, created by Pereira et al., was utilized to dock 2,950 ligands with 40 receptors. When tested with 95,000 decoy molecules against these receptors, it performed exceptionally well [14]. Using a multi-objective automated replacement technique, another method assessed the cyclin-dependent kinase-2 inhibitor's physicochemical characteristics, biochemical activity, and form similarity in order to maximize its potency profile [15].

In addition to being used to find possible drug candidates, quantitative structure-activity relationship modelling tools have evolved into AI-based techniques including decision trees, random forests, support vector machines, and linear discriminant analysis (LDA), which speed up QSAR analysis [16-18]. King et al. Reported that six AI algorithms ranked anonymous compounds by biological activity with nearly the same accuracy as traditional methods [19].

AI has transformed drug research and discovery in many ways. Some of its major contributions are:

Visual screening:

AI allows efficient screening of large chemical libraries to find potential drug candidates likely to bind a specific target. By predicting interactions and binding affinities, it helps researchers prioritize compounds for testing, saving time and resources.

De novo drug design:

With reinforcement learning and generative models, AI can design new drug-like molecules. By learning from chemical libraries and experimental data, it expands chemical space and supports the discovery of innovative drug candidates.

Target identification:

AI can analyse genetic, proteomic, and clinical data to identify therapeutic targets. By revealing disease-related targets and pathways, it helps guide the design of drugs that modulate biological processes.

Optimization of drug candidates:

AI can analyse and optimize drug candidates by assessing efficacy, safety, and pharmacokinetics, helping researchers improve effectiveness while reducing side effects.

Virtual screening:

AI enables rapid screening of large chemical libraries to identify candidates likely to bind a target. By predicting interactions and binding strength, it helps prioritize compounds for testing, saving time and resources.

Structural activity relationship (SAR) modelling:

AI models can link chemical structures to biological activity, helping researchers design drug candidates with desirable traits like high potency, selectivity, and good pharmacokinetics.

Drug repurposing:

AI can analyse large biomedical datasets to find existing drugs with potential for new diseases. Repurposing approved drugs in this way speeds up discovery and lowers costs.

Toxicity prediction:

AI can predict drug toxicity by analysing chemical structures and properties.

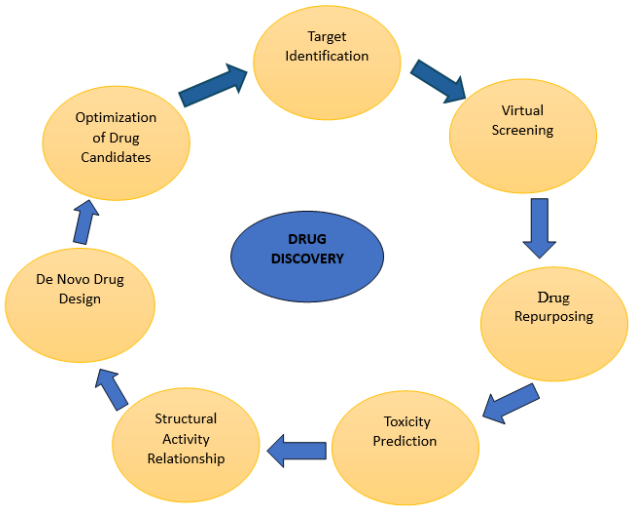


Figure 1: Artificial Intelligence in Drug Discovery

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.

ML models trained on toxicology data can identify harmful effects or risky structural features, helping researchers focus on safer compounds and reduce adverse outcomes in trials. Overall, AI-driven methods in drug research can speed up the identification, optimization, and design of new therapies, leading to more efficient and effective medicines[20].

Application of AI in drug discovery:

1. Drug screening and lead discovery:

In silico techniques and virtual screening driven by AI have significantly enhanced the search for possible lead compounds in drug discovery. Unlike traditional high-throughput screening, these methods reduce expenses and speed up the process by rapidly analysing massive chemical libraries using computational tools [21,22]. The algorithms used in machine learning (ML) are essential to these techniques. Quantitative structure–activity relationship (QSAR) models, for instance, are constructed using them to forecast a compound's biological activity based on its chemical structure [23]. The compounds that are most likely to bind to the target of interest are highlighted by these models, which also aid in screening enormous chemical libraries [24].

2. Drug optimization and design:

Through the improvement of crucial characteristics like solubility, stability, and bioavailability, AI-driven approaches are revolutionizing the medication development process[25]. Machine learning algorithms are capable of reliably predicting crucial parameters by analysing vast databases of chemical structures and properties [26]. For example, 1,000–5,000 data points are frequently used in QSAR models for water solubility [27]. On the other hand, drug stability can be predicted using deep learning (DL) models in a variety of scenarios [28]. Open databases like the UniProt Consortium can be used by researchers to anticipate the function of proteins [29]. To gather information on protein sequences from many species, as well as the Protein Data Bank (PDB) [30]. Following preprocessing, cleaning, and batch downloading, this data can be utilized to train prediction models [31]. Additionally, DL methods, such generative adversarial networks (GANs), can expand the chemical space in drug discovery by producing novel chemical compounds with desirable features [32].

3. Preclinical and clinical development:

For effective medication development, pharmacokinetics (PK), pharmacodynamics (PD), and toxicity must be predicted. These forecasts have been significantly enhanced by AI, particularly machine learning (ML). In order to estimate important parameters like drug interactions and absorption, distribution, metabolism, and excretion (ADME), machine learning (ML) can process enormous datasets of chemical compounds and their PK/PD properties. For example, some models can simulate the distribution of medications in the body, whereas deep learning (DL) models can forecast how they are absorbed across membranes [33]. Additionally, AI can forecast toxicity profiles and drug reactions. ML algorithms assist researchers choose safer drug options and lower the likelihood of side effects by identifying patterns between chemical structures and toxicity results [34,35].

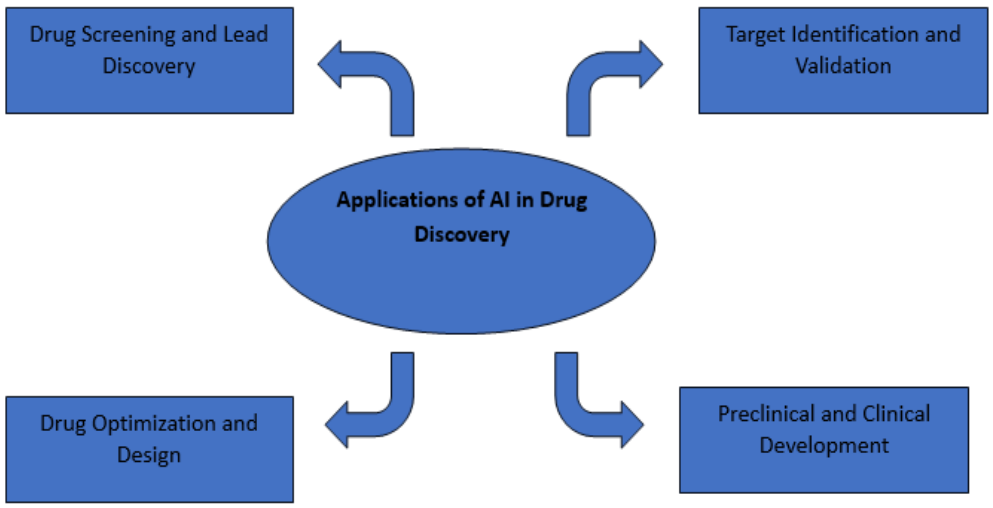


Figure 2: Applications of Artificial Intelligence in Drug Discovery. The various AI approaches are Drug screening and lead discovery, Target identification, preclinical and clinical development etc.

AI has revolutionized patient recruiting, data analysis, and clinical trial design, increasing the efficiency of research. AI can optimize study design by choosing objectives, determining sample size, and finding appropriate patient groups by evaluating data from previous trials [36]. Platforms with AI capabilities enhance patient recruitment by identifying possible participants and interacting with them through tailored outreach and targeted advertisements [22]. AI is crucial to real-time data analysis and monitoring. ML algorithms are able to continuously monitor trial data in order to identify adverse events, detect safety signals, and assess the efficacy of treatments [37]. This helps researchers make informed changes, such as adjusting doses or adding new treatment arms, resulting in faster and more efficient clinical trials fig. 2.

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.

AI techniques in drug discovery:

I. Machine learning algorithms using supervised and unsupervised learning:

The two primary categories of these algorithms are supervised and unsupervised. The system is taught using labelled data in supervised learning, which is particularly helpful for forecasting the efficacy of medications and potential adverse effects. For instance, models can assist in forecasting the potential effectiveness of novel compounds when they are trained on datasets of compounds that have demonstrated action against a target [38]. Similarly, by utilising historical data on bad drug reactions, supervised learning can forecast side effects [39]. Unsupervised learning, on the other hand, uses unlabelled data to uncover connections and hidden patterns. By clustering chemicals with similar chemical structures or genes with similar expression patterns, it can be utilised to find new therapeutic targets. Random forests, which can handle big, complicated datasets and emphasise the most crucial aspects, are one of the frequently employed machine learning techniques in drug development [40]. SVM is effective for classification tasks and for finding the best decision boundaries [41]. In order to create a powerful prediction model, gradient boosting machines integrate numerous weak models [42]. By facilitating the quick analysis of big datasets and locating interesting drug candidates, these algorithms play a crucial role in speeding up drug discovery[43].

II. Deep learning utilizing neural networks:

In drug design, deep neural networks (DNNs), particularly CNNs and RNNs, are effective tools. CNNs can predict protein–ligand binding affinities and other important qualities by learning intricate patterns. They are very good at processing image-like data, such as 2D or 3D molecular structures [44]. Sequential data, such as protein sequences or chemical reaction pathways, are ideally suited for RNNs. They are able to examine enzymes that break down drugs, forecast protein-protein interactions, and even create novel compounds that resemble drugs [45].

A crucial stage in drug design, structure–activity relationship (SAR) modelling, is being transformed by deep learning (DL). DL can automatically discover intricate patterns from molecular data, providing more accurate and dependable predictions than typical SAR models that depend on manually selected characteristics. CNNs, for instance, are capable of capturing intricate chemical information in order to forecast characteristics like metabolic stability, solubility, and binding affinity [46].

Another popular application of deep learning (DL) is the prediction of drug toxicity. Large datasets of toxicity profiles and chemical structures are analysed by DL models to find trends that people would overlook. By enabling the early detection of harmful substances, this reduces the possibility of unfavourable trial outcomes and enhances medication safety [47, 48].

III. Natural language processing in identification of novel candidates:

Finding new drug candidates and biomarkers involves analysing vast amounts of unstructured text from scientific publications, patents, and electronic health records (EHRs) with the aid of natural language processing (NLP). NLP can extract important data to find possible targets and treatments, including chemical structures, biological processes, and experimental results. It can also evaluate intellectual property, identify new chemicals, and point forth joint venture or licensing prospects by analysing patents [49].

Transformer models have advanced NLP with their self-attention mechanism, which captures long-range dependencies in data. Unlike RNNs, they process sequences in parallel, improving performance in tasks such as translation, text generation, and speech recognition. Successful models like BERT, GPT, and T5 have also inspired applications in computational biology, especially in Protein Language Models (PLMs) [50].

Using transformer architectures, protein language models (PLMs) interpret biological sequences by interpreting proteins as “text” composed of “alphabets” of amino acids. PLMs can predict structures (like AlphaFold2), identify functions, create proteins, and evaluate the effects of mutations after being trained on big protein datasets. Transfer learning with pretrained models like ESM-2 and ProtGPT2, unsupervised learning from resources like UniProt, and geometric awareness with models like Omega Fold are some of their benefits [51]. These models use 3D structural information. Future directions include energy-based optimization for de novo protein creation and multimodal protein models, despite obstacles such high computational cost and restricted interpretability [52, 53]. The ESM-2 model from Meta, which has 15 billion parameters and can predict protein structures from sequences, is an example of how powerful transformers are at deciphering the biological "language” [54].

NLP can also analyse EHRs to find patterns linking patient characteristics, disease progression, and treatment outcomes. This supports the discovery of new biomarkers, prediction of therapy responses, and development of personalized treatment plans [55].

IV. Reinforcement learning and evolutionary algorithms:

Drug optimization can be effectively approached by using reinforcement learning (RL), which views the problem as a decision-making one. In this case, the RL agent learns to change chemical structures or dosages in order to minimize toxicity and increase benefits like efficacy. In order to discover the best answers that might not be immediately apparent to researchers, RL can comb through vast amounts of medication candidates and therapy approaches [56, 57].

Natural selection serves as the inspiration for evolutionary algorithms, which are useful for molecular structure optimization. They produce and hone molecules with

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.

desired qualities using procedures like selection and mutation. Iteratively altering molecules and evaluating their fitness (such as binding affinity or drug-likeness) allows these algorithms to effectively search chemical space and find new, improved therapeutic candidates [58].

V. Generative AI:

ChatGPT and other generative AI tools have enormous promise for clinical practice and healthcare education. In order to improve patient care, they can assist with duties like prescription review, drug interaction checks, and adverse reaction monitoring in pharmacies. They are still not widely used in pharmacy education, though, and further study is required to overcome implementation issues [59]. Beyond clinical duties, generative AI may also build massive datasets to train predictive models, expanding its importance in medical research. With two networks—a generator that creates synthetic data and a discriminator that assesses it—GANs are a potent deep learning framework among these methods. Through iterative training, the generator develops until it provides very realistic data, supporting applications such as medical imaging, super-resolution, and data augmentation [60]. For instance, Super-Resolution GAN has demonstrated efficacy in enhancing low-resolution images, rendering it valuable for video surveillance and medical diagnostics [61]. Moving forward, unlocking the full potential of LLM-driven biotechnology will require clear performance benchmarks, improved model transparency, and stronger collaboration between computational and life science fields. This convergence is set to transform research methods and industrial processes across the biological sciences [62]. Fig .3

VI. Molecular fingerprinting:

A computer technique called molecular fingerprinting uses basic numerical representations of chemical structures, like feature arrays or bit vectors, to convey significant physicochemical and structural characteristics. In materials research and drug development, these fingerprints are extremely helpful since they facilitate quick comparison, similarity searching, and machine learning applications. Common varieties include path-based fingerprinting for bond sequences, circular fingerprints such as Extended Connectivity fingerprinting (ECFP) for substructure identification, and 3D fingerprints for molecular shape capture[63]. Molecular fingerprinting encodes chemical "grammar," such as functional groups and bonds, as opposed to verbal semantics, which are processed by natural language processing (NLP). For applications like virtual screening and toxicity prediction, classical fingerprinting uses cheminformatics methods, whereas NLP models such as transformers may analyse molecules using SMILES strings. Instruments like RDKit [64] and A popular tool for effectively creating fingerprints is ChemDes, which aids in bridging the gap between chemistry and artificial intelligence [65].

Nevertheless, there are still issues with enhancing fingerprint techniques for various molecular datasets and raising their prediction precision for biological applications. To better find active compounds in large-scale screenings, for instance, ECFPs still require improvement, notwithstanding their effectiveness in structure–activity modelling and similarity searches [66].

VII. Graph neural networks:

Graph Neural Networks (GNNs) are advanced deep learning models built to cope with graph-structured data, which is made up of nodes (entities) and edges (relationships). Unlike standard neural networks that handle grid-like input such as images or text, GNNs may capture intricate connections in non-Euclidean regions. This makes them particularly valuable for applications like social network research, drug development, and recommendation systems. GNNs work through a message-passing process, where nodes acquire and integrate information from their neighbours to learn effective representations of both local and global graph structures. Popular GNN architectures include Graph Convolutional Networks (GCNs), which use spectral convolution; Graph Attention Networks (GATs), which apply attention to highlight key neighbours; and GraphSAGE, which provides scalable inductive learning [67,68]. Although GNNs have gained success in many sectors, they still face obstacles like as handling very large graphs and the issue of over-smoothing in deeper models. Future research is anticipated to focus on self-supervised learning methods and enhancing model explainability, especially for sensitive domains like healthcare. As the use of graph-structured data rises, GNNs are projected to dramatically advance machine learning across diverse areas [69].

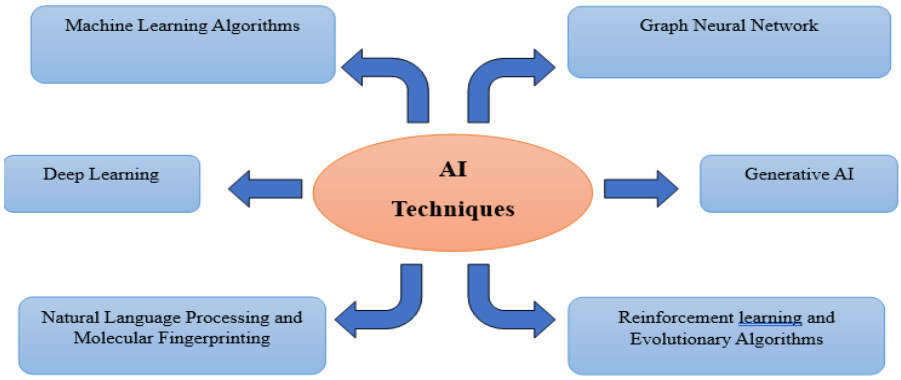


Figure 3: Artificial Intelligence Techniques in Drug Discovery

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.

Current challenges and limitations:

a. Interpretability and transparency:

The complexity and lack of transparency of AI systems pose significant challenges to their adoption. Many models, particularly deep neural networks (DNNs), operate as "black boxes," making it difficult to understand how they make judgments [70]. The lack of interpretability and openness raises questions about trust, accountability, and possible prejudice. In healthcare, for example, understanding how an AI system gets a diagnosis is critical for physicians to make educated decisions and ensure patient safety [71].

b. Data quality and availability:

The scarcity of high-quality annotated datasets for training is a significant barrier to AI-powered drug discovery. Data are frequently gathered from a variety of sources, including chemical structures, biological testing, and clinical trials, resulting in heterogeneity. Combining and standardizing various datasets into a single format for AI training is both difficult and time-consuming [72]. Biases in training data can drastically degrade model performance and dependability. For example, if a dataset mostly represents one demographic or disease group, the model may develop biases that limit its accuracy and utility in real-world scenarios [73,74]. To address these challenges, careful data curation, strong preprocessing methods, and strategies to reduce bias and improve data representativeness are needed.

c. Ethical challenges of using AI:

While there are many potential benefits to using artificial intelligence (AI) in pharmacy practice, there are also significant ethical issues. In the Middle East and North Africa (MENA) area, pharmacy professionals recently participated in a survey that revealed important concerns such cybersecurity threats (58.9%), patient data protection (58.9%), potential job loss (62.9%), and unclear legal regulations (67.0%) [75]. These results emphasise how crucial it is to adopt AI with beneficence, justice, openness, and informed consent. Greater tech awareness was associated with higher ethical concern scores, highlighting the need for ethical education and guidelines. Significant ethical issues also surround patient safety and clinical trials. Giving informed permission and being fully aware of the hazards are essential to the integrity of trials. In addition to violating people's rights, this procedure puts patient safety at risk. Even while the FDA and other regulatory agencies enforce protections, problems like the exploitation of vulnerable people continue to exist. Responsible AI usage in medicine necessitates open frameworks, interdisciplinary cooperation, and a strong dedication to patient care [76]. AI adoption in pharma faces major ethical and legal challenges, including algorithmic bias, lack of transparency in decisions, data privacy risks, and unclear liability. Regulatory bodies are still adjusting to oversee AI use, highlighting the need for standardized guidelines that protect patients while supporting innovation. As AI advances, it is crucial to balance its potential benefits with strong ethical governance to ensure safe drug development, better patient care, and sustained public trust in pharmaceutical innovation [77].

d. Integration into existing drug development process:

Due to the rigid procedures and tried-and-true methodologies used in traditional pharmaceutical workflows, integrating AI tools into drug development presents significant obstacles [36]. Significant adjustments to workflows, infrastructure, and expertise may be necessary for the adoption of AI. Concerns around intellectual property, data privacy, and the effect on pharmaceutical sector employment may also impede adoption [78].

CONCLUSION

The process of finding new drugs is still difficult and expensive, but developments in screening technology, molecular design, and computational modelling are revolutionizing early-stage research. These developments save time and money by making it possible to discover possible medication candidates more quickly and accurately. But in order to guarantee trustworthy results, problems like data constraints and regulatory obstacles need to be resolved. For these advancements to be translated into safer and more effective treatments, researchers, industry, and politicians must work together more closely.

Source of Support: The author(s) received no financial support for the research, authorship, and/or publication of this article

Conflict of Interest: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

1. Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. Ann R Coll Surg Engl. 2004;86(5):334-338.

2. Miles JC, Walker AJ. The potential application of artificial intelligence in transport. IEE Proc Intell Transp Syst. 2006;153(3):183-198.

3. Yang Y, Siau K. Qualitative research on marketing and sales in the artificial intelligence age. In: Proceedings of the 13th Annual Midwest Association for Information Systems Conference; 2018 May 17-18; St Louis, MO. 2018.

4. Wirtz BW, Weyerer JC, Geyer C. Artificial intelligence and the public sector—applications and challenges. Int J Public Adm. 2019; 42:596-615.

5. Smith RG, Farquhar A. The road ahead for knowledge management: an AI perspective. AI Mag. 2000;21(4):17-20.

6. Lamberti MJ, Wilkinson M, Donzanti BA, Wohlhieter GE, Parikh S, Wilkins RG, et al. A study on the application and use of artificial intelligence to support drug development. Clin Ther. 2019;41(8):1414-1426.

7. Vyas M, Thakur S, Riyaz B, Bansal KK, Tomar B, Mishra V. Artificial intelligence: the beginning of a new era in pharmacy profession. Asian J Pharm. 2018; 12:72-76.

8. Ciallella HL, Zhu H. Advancing computational toxicology in the big data era by artificial intelligence: data-driven and mechanism-

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.

driven modeling for chemical toxicity. Chem Res Toxicol. 2019; 32:536-547.

9. Chan HCS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing drug discovery via artificial intelligence. Trends Pharmacol Sci. 2019;40(8):592-604.

10. Brown N. In Silico Medicinal Chemistry: Computational Methods to Support Drug Design. Cambridge: Royal Society of Chemistry; 2016.

11. Firth NC, Atrash B, Brown N, Blagg J. MOARF, an integrated workflow for multi objective optimization: implementation, synthesis, and biological evaluation. J Chem Inf Model. 2015;55(6):1169-1180.

12. Pereira JC, Caffarena ER, Dos Santos CN. Boosting docking-based virtual screening with deep learning. J Chem Inf Model. 2016;56(12):2495-2506.

13. Sellwood MA, Ahmed M, Segler MHS, Brown N. Artificial intelligence in drug discovery. Future Med Chem. 2018;10(17):2025-2028.

14. Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects. Drug Discov Today. 2019;24(3):773-780.

15. Zhu H. Big data and artificial intelligence modeling for drug discovery. Annu Rev Pharmacol Toxicol. 2020; 60:573-589.

16. Zhang L, Tan J, Han D, Zhu H. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug Discov Today. 2017;22(11):1680-1685.

17. Jain N, Gupta S, Sapre N, Sapre NS. In silico de novo design of novel NNRTIs: a bio-molecular modelling approach. RSC Adv. 2015;5:14814-14827.

18. Wang Y, et al. A comparative study of family-specific protein-ligand complex affinity prediction based on random forest approach. J Compute Aided Mol Des. 2015;29(4):349-360.

19. King RD. Comparison of artificial intelligence methods for modeling pharmaceutical QSARs. Appl Artif Intell. 1995; 9:213-233.

20. Shah H, Chavda V, Soniwala MM. Applications of bioinformatics tools in medicinal biology and biotechnology. In: Bioinformatics Tools for Pharmaceutical Drug Product Development. 2023. p.95-116.

21. Naithani U, Guleria V. Integrative computational approaches for discovery and evaluation of lead compound for drug design. Front Drug Discov. 2024.

22. Paul D, et al. Artificial intelligence in drug discovery and development. Drug Discov Today. 2021;26(1):80-93.

23. Ching T, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface. 2018;15(141):20170387.

24. Niazi SK, Mariam Z. Recent advances in machine-learning-based cheminformatics: a comprehensive review. Int J Mol Sci. 2023;24(14).

25. Vidhya KS, et al. Artificial intelligence's impact on drug discovery and development from bench to bedside. Cureus. 2023;15(10).

26. Singh S, et al. Artificial intelligence and machine learning in pharmacological research: bridging the gap between data and drug discovery. Cureus. 2023;15(8).

27. Piir G, et al. Best practices for QSAR model reporting: physical and chemical properties, ecotoxicity, environmental fate, human health, and toxicokinetic endpoints. Environ Health Perspect. 2018;126(12):126001.

28. An F, et al. Machine learning model for prediction of drug solubility in supercritical solvent: modeling and experimental validation. J Mol Liq. 2022; 363:119901.

29. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res. 2025;53(D1): D609-D617.

30. Berman HM, Burley SK. Protein Data Bank (PDB): fifty-three years young and having a transformative impact on science and society. Q Rev Biophys. 2025;58: e9.

31. Chen JY, et al. Evaluating the advancements in protein language models for encoding strategies in protein function prediction: a comprehensive review. Front Bioeng Biotechnol. 2025;13.

32. Sousa T, et al. Generative deep learning for targeted compound design. J Chem Inf Model. 2021;61(11).

33. Wu K, et al. Predicting pharmacodynamic effects through early drug discovery with artificial intelligence-physiologically based pharmacokinetic (AI-PBPK) modelling. Front Pharmacol. 2024;15.

34. Tran TTV, et al. Artificial intelligence in drug toxicity prediction: recent advances, challenges, and future perspectives. J Chem Inf Model. 2023;63(9).

35. Yang S, Kar S. Application of artificial intelligence and machine learning in early detection of adverse drug reactions (ADRs) and drug-induced toxicity. Artif Intell Chem. 2023;1(2).

36. Bretz F, et al. Adaptive designs for confirmatory clinical trials. Stat Med. 2009;28(8).

37. Chow SC, Liu JP. Design and Analysis of Clinical Trials: Concepts and Methodologies. 2nd ed. Hoboken (NJ): Wiley; 2008.

38. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463-477.

39. Toni E, et al. Machine learning techniques for predicting drug-related side effects: a scoping review. Pharmaceuticals (Basel). 2024;17(6).

40. Breiman L. Random forests. Mach Learn. 2001;45(1):5-32.

41. Awad M, Khanna R. Support vector machines for classification. In: Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Berkeley (CA): Apress; 2015. p.39-66.

42. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001;29(5):1189-1232.

43. Patel L, Shukla T, Huang X, Ussery DW, Wang S. Machine learning methods in drug discovery. Molecules. 2020;25(22):5277.

44. Alzubaidi L, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021;8(1).

45. Grisoni F, et al. Bidirectional molecule generation with recurrent neural networks. J Chem Inf Model. 2020;60(3).

46. Visan AI, et al. Integrating artificial intelligence for drug discovery in the context of revolutionizing drug delivery. Life (Basel). 2024;14(2).

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.

47. Guo W, Liu J, Dong F, Song M, Li Z, Khan MKH, et al. Review of machine learning and deep learning models for toxicity prediction. Exp Biol Med (Maywood). 2023;248(21).

48. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: toxicity prediction using deep learning. Front Environ Sci. 2016; 3:80.

49. Kosonocky CW, Wilke CO, Marcotte EM, Ellington AD. Mining patents with large language models elucidates the chemical function landscape. Digital Discovery. 2024;3(6).

50. Zhang S, et al. Applications of transformer-based language models in bioinformatics: a survey. Bioinform Adv. 2023;3(1).

51. Wu R, et al. High-resolution de novo structure prediction from primary sequence. bioRxiv. 2022. doi:10.1101/2022.07.21.

52. Bepler T, Berger B. Learning the protein language: evolution, structure, and function. Cell Syst. 2021;12(6).

53. Ferruz N, Schmidt S, Höcker B. ProtGPT2 is a deep unsupervised language model for protein design. Nat Commun. 2022;13(1):4348.

54. Lin Z, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science. 2023;379(6637).

55. Cowie MR. Electronic health records to facilitate clinical research. Clin Res Cardiol. 2017;106(1):1-9.

56. Popova M, et al. Deep reinforcement learning for de novo drug design. Sci Adv. 2018;4(7): eaap7885.

57. Jayaraman P, et al. A primer on reinforcement learning in medicine for clinicians. NPJ Digit Med. 2024;7(1).

58. Lameijer EW, et al. Evolutionary algorithms in drug design. Nat Comput. 2005;4(3).

59. Mortlock RC, et al. Generative artificial intelligence (Gen-AI) in pharmacy education: utilization and implications for academic integrity: a scoping review. Explore Res Clin Soc Pharm. 2024;15.

60. Elahi M, et al. A comprehensive literature review of the applications of AI techniques through the lifecycle of industrial equipment. Discov Artif Intell. 2023;3(1).

61. Lan L, et al. Generative adversarial networks and its applications in biomedical informatics. Front Public Health. 2020;8.

62. Li W, et al. large language model for knowledge synthesis and AI-enhanced biomanufacturing. Trends Biotechnol. 2025.

63. Li Z, et al. Fingerprinting interactions between proteins and ligands for facilitating machine learning in drug discovery. Biomolecules. 2024;14(1).

64. Landrum GA. RDKit: open-source cheminformatics. 2013.

65. Dong J, et al. ChemDes: an integrated web-based platform for molecular descriptor and fingerprint computation. J Cheminform. 2015;7.

66. Ucak UV, et al. Reconstruction of lossless molecular representations from fingerprints. J Cheminform. 2023;15(1).

67. Zaman AM, et al. The leading global health challenges in the artificial intelligence era. Front Public Health. 2023;11.

68. Wu Z, et al. A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn Syst. 2020;32(1).

69. Corso G, et al. Graph neural networks. Nat Rev Methods Primers. 2024;4(1).

70. Campolo C, et al. Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. Front Hum Dyn. 2024. doi:10.3389/fhumd.2024.1421273.

71. Kiseleva A, et al. Transparency of AI in healthcare as a multilayered system of accountabilities: between legal requirements. Front Artif Intell. 2022;5.

72. Blanco-González A, et al. The role of AI in drug discovery: challenges, opportunities, and strategies. Pharmaceuticals (Basel). 2023;16(6).

73. Kokudeva M, et al. Artificial intelligence as a tool in drug discovery and development. World J Exp Med. 2024;14(3).

74. Gichoya JW, et al. AI pitfalls and what not to do: mitigating bias in AI. Br J Radiol. 2023;96(1150).

75. Hasan HE, et al. Ethical considerations and concerns in the implementation of AI in pharmacy practice: a cross-sectional study. BMC Med Ethics. 2024;25(1).

76. Boudi AL, et al. Ethical challenges of artificial intelligence in medicine. Cureus. 2023;16(11).

77. Shaki F, et al. Artificial intelligence in pharmaceuticals: exploring applications and legal challenges. Pharm Biomed Res. 2024;10(1):1-10.

78. Dwivedi YK, et al. Artificial Intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. Int J Inf Manage. 2021;57.

For any questions related to this article, please reach us at: globalresearchonline@rediffmail.com

New manuscripts for publication can be submitted at: submit@globalresearchonline.net and submit_ijpsrr@rediffmail.com

©Copyright protected. Unauthorised republication, reproduction, distribution, dissemination and copying of this document in whole or in part is strictly prohibited.