



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



Review Article

A Review Of Artificial Intelligence In Drug Discovery And Development

Somnath S. Davkhar*, Devadhe Sanika S., Khedkar Namdev Bhagwan

*Nav Maharashtra Shikshan Mandal Abasaheb Kakade College Of B. Pharmacy Bodhegaon , Tal. Shevgaon.
Dist. Ahmednagar*

ARTICLE INFO

Received: 25 Jan 2024

Accepted: 29 Jan 2024

Published: 03 Mar 2024

Keywords:

Artificial Intelligence (AI)
Technologies,

DOI:

10.5281/zenodo.10603065

ABSTRACT

From initial molecule discovery to bringing a new drug to market, it takes at least 10 years to complete the process of making a drug, costing approximately 2.5 billion dollars. Over a past decade, Artificial Intelligence (AI) have brought ease to the process of developing a drug and helping companies save time and money. AI tools are revolutionizing nearly every stage of the drug discovery process, offering substantial potential to reshape the speed and economics of the industry. AI has recently started to gear up its application in various pharmaceutical sectors viz., drug discovery, drug repurposing, clinical trials, pharmaceutical productivity, virtual screening, drug design, etc. In this survey as a major part, we have discussed on drug discovery and development which have transformed thought "developing treatments can be a risky business". We then discourse on reinforcement and machine learning. We also have some crosstalk on software tools and techniques of AI used in drug development. This review article highlights enforcing AI utilization in pharmaceutical industry. We expect that this survey provides a panoramic review on AI in drug discovery and development.

INTRODUCTION

The process of drug development can be defined in five steps, but the time taken to cover these steps in the real world spans a decade, and sometimes more. The journey of new medicines from a laboratory to a pharmacy store is lengthy and complex, with costs running into billions of Dollars. Moreover, the success rate of launching a drug to market from phase 1 clinical trials is daunting, less than 10%. The COVID-19

pandemic has emphasized the need for novel drug discovery process. However, the journey from conceptualizing a drug to its eventual implementation in clinical settings is a long, complex, and expensive process, with many potential points of failure. Over the past decade, a vast growth in medical information has coincided with advances in computational hardware (cloud computing, GPUs, and TPUs) and the rise of deep learning. Living in the era of big data, the real fight

***Corresponding Author:** Somnath S. Davkhar

Address: Nav Maharashtra Shikshan Mandal Abasaheb Kakade College Of B. Pharmacy Bodhegaon , Tal. Shevgaon. Dist. Ahmednagar

Email ✉: nkhedkar2002@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.



for drug discovery is how to better utilize small available data to predict outside of the known space into dark space. Artificial intelligence (AI) is rapidly transforming the pharmaceutical industry, with the potential to revolutionize drug discovery and development. The pharmaceutical industry is embracing the power of AI and machine learning (ML) to transform the way medicines are discovered, developed and administered. These technologies are revolutionizing various areas, including drug research, personalized medicines, clinical trials, drug safety, supply chain management and regulatory compliance. By harnessing the compatibilities of AI and ML, pharmaceutical companies are achieving faster results, improving patient outcomes, and enhancing efficiency in the industry. Popular applications of AI in drug discovery includes virtual screening, retrosynthesis, reaction prediction, prediction of the properties of potential drug candidate's such as their toxicity and efficacy. The use of AI in drug discovery is still in its early stages, but the potential is enormous. In this survey, we mainly focus on AI-driven drug discovery and it's commercialisation. We have provided the

application and techniques involved in discovery of small molecule drugs. Firstly, we discussed an overview on key applications in drug discovery of AI and point out a collection of previously published perspectives, reviews, and surveys. We present applications of AI at various stages of drug discovery pipelines, including the inherently computational approaches of the de novo design and prediction of a drugs likely properties. Also. Know contemporary AI methods such as graph neural networks, reinforcement learning and generated models, along with structure-based methods, that can contribute to drug discovery applications and analysis of drug responses is also explored. With knowledge on data and related techniques, including model architecture and machine learning, will be elaborated. Finally, recent developments and investments in AI-based start-up companies for biotechnology, drug design and their current progress, hopes and promotion are discussed in this article. By gathering informative data with the several surveyed papers, we expect this survey not only provides a comprehensive overview of AI I drug discovery but also serve as a learning resource for researchers intended in this interdisciplinary field.

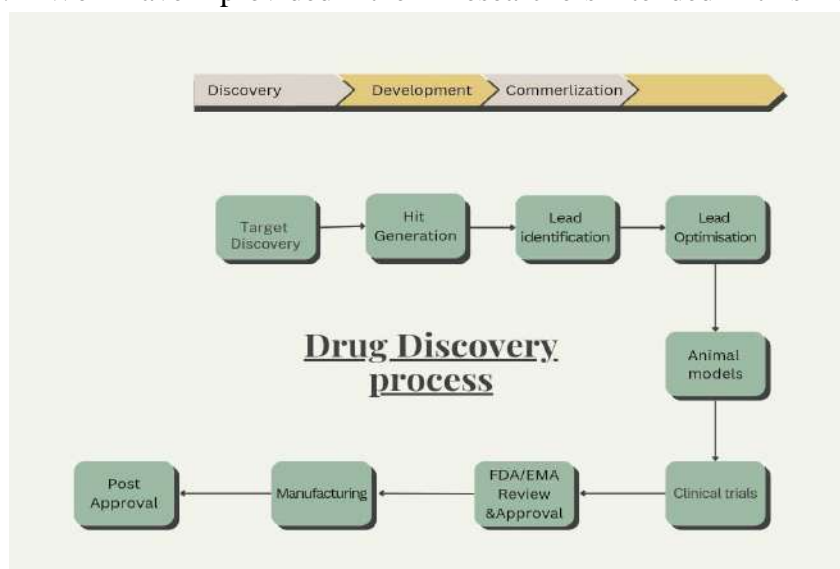


Fig.no.1 Drug Discovery process

**Discovery
Target discovery**

Machine learning algorithms, especially deep learning methodologies, have gained attention and shown promising results in pharmaceutical fields. Deep learning, also known as deep neural networks, involves multiple hidden layers of nodes for data processing and feature extraction. Recent deep learning architectures like GANs, recurrent neural networks, and transfer learning techniques have been applied in healthcare, including small-molecule design and aging research. These advanced deep learning methods have outperformed traditional machine learning approaches in healthcare applications. The use of deep learning in pharmaceutical areas has the potential to revolutionize drug discovery and improve patient care. Deep learning techniques have been used to predict pharmacological effects of drugs based on transcriptional data of drug-perturbed cell lines. Pun et al. utilized bioinformatic and deep learning models trained on disease-specific multiomic and text-based data to identify potential therapeutic targets for amyotrophic lateral sclerosis (ALS), resulting in the discovery of 18 drugable genes. Fabris et al. developed a deep learning-based method that

incorporated gene and protein features to identify human genes associated with multiple age-related diseases. West et al. used a deep learning ensemble trained on transcriptomic profiles of embryonic and adult cells to uncover a novel target (COX7A1) involved in the embryonic-fetal transition, providing insights into our understanding of normal development, epimorphic tissue regeneration, and cancer. Large language models like BioGPT and ChatPandaGPT aid in therapeutic target discovery through rapid biomedical text mining, connecting diseases, genes, and biological processes for identifying mechanisms, drug targets, and biomarkers. These models understand natural language and complex scientific concepts, making them valuable tools for accelerating disease hypothesis generation. However, they may perpetuate human biases and preconceived notions as they are trained on human-generated text. Additionally, their reliance on published data may limit their ability to identify genuinely novel targets. To ensure the discovery of truly novel and pertinent targets, it is important to acknowledge these limitations and complement their use with other models.

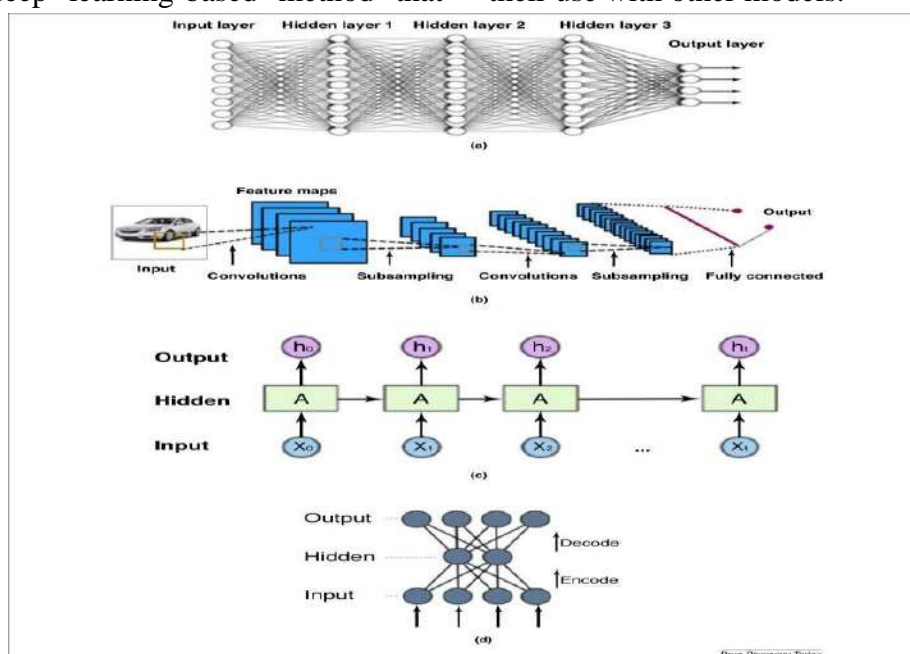


Fig no. 2 Lead identification and optimization

The explosion of biomedical data in recent years has created challenges for data analysis, but AI can help tackle these challenges. AI algorithms have the advantage of processing complex biomedical networks of data, revealing patterns and relationships that may not be apparent to humans. The use of AI in biomedical research can lead to a better understanding and treatment of diseases. AI has made notable contributions in facilitating biomarker and target identification. AI has the potential to revolutionize the field of biomedical research by providing insights and solutions that were previously unattainable. Target identification is crucial in drug discovery, but it remains challenging despite advancements in experimental and omic technologies. The integration of multiomic data with AI algorithms is a promising approach for target identification. This paper focuses on the application of AI algorithms to target identification and aims to encourage the integration of AI technologies into drug discovery pipelines. The emergence of the AI-driven drug discovery era offers a progressive outlook for the field. AI-driven target identification has the potential to revolutionize the drug discovery process and improve therapeutic outcomes. AI/ML methods in drug discovery are becoming more advanced and will have a significant impact on various aspects of the process, such as lead finding and optimization. These methods often use ML-models to predict the properties of small molecules based on their 2D chemical structures. However, the lack of data, especially for new targets, makes it challenging to build effective ML-models for structure-activity relationships. The BIOVIA Generative Therapeutics Design (GTD) application utilizes 3D structural models of ligand protein interactions, allowing for a more accurate representation of desired features. Using a dataset related to the discovery of SYK inhibitors, the GTD application effectively addresses common problems in lead finding and optimization. ML-

models, when used alongside calculable constraints like MW, HBAs, HBDs, and aromatic rings, can effectively shape the chemical space for a specific design problem. However, early ML-model based lead optimization projects often suffer from poor models of biological activity. In contrast, the traditional structure-based design approach relies on analyzing ligand protein interactions in crystal structures to measure target engagement. The combination of observed ligand protein interactions and chemical design space criteria has been central to structure-based design since its inception. The quality of ML-models greatly impacts the outcome of the lead optimization process, which focuses on enhancing biological activity. In cases where highly predictive ML-models for biological activity are lacking, incorporating 3D information of ligand protein interactions becomes crucial in the generative iteration cycles. The Generative Therapeutics Design (GTD) method, developed by Honeycutt et al. in 2021, utilizes these interactions as optimization criteria through simple pharmacophore features. These features can be implemented as binary constraints or quantitative parameters to be optimized against, such as fit values, shape similarity, or docking scores. GTD offers a new approach to molecule optimization by integrating ligand protein interactions into the optimization process, improving the potential.

- **Development**

- **Clinical trials design and monitoring:**

Clinical trials are a crucial and expensive part of drug development, with a failed trial resulting in a loss of hundreds of millions to billions of dollars. High trial failure rates are often due to suboptimal patient selection and monitoring techniques. Recent advances in artificial intelligence (AI) can help improve the design of clinical trials and increase success rates. AI solutions can enhance hypothesis generation, drug discovery, cohort composition, monitoring, adherence, and endpoint



selection. Collaboration is needed to establish common protocols for data collection and organization to reduce errors in AI output. Integration of AI in clinical trials shows promise for more successful and cost-effective drug development. However, there is a lack of specific

regulatory guidance on the use of AI in clinical trials. Artificial intelligence (AI) can improve trial design and increase success rates. Moreover, there is a need for collaboration and regulatory guidance to ensure ethical and safe implementation of AI in clinical trials.



Fig.no. 3 Benefits of leveraging AI in clinical trials

1. AI can improve patient recruitment and monitoring in clinical trials, reduce dropout rates, and enhance data quality. It can analyze various data types, predict adverse events, and simplify data collection using wearable technologies. AI also helps researchers uncover meaningful insights from the massive amount of data generated in clinical trials.
2. AI can improve patient recruitment by identifying and screening potential participants based on inclusion and exclusion criteria, reducing time and cost.
3. AI can improve patient monitoring during clinical trials by detecting and predicting adverse events, leading to better patient safety and protocol adjustments.
4. Wearable devices and sensors combined with AI can collect real-time data to identify patterns indicating potential adverse events or complications.
5. AI can improve subject retention by identifying factors associated with high dropout risk and predicting the probability of dropout, allowing proactive outreach.
6. AI can minimize the effects of confounders and improve data quality by combining with wearable technologies for simplified data collection.
7. AI can analyze data collected by wearable technologies and diagnostic devices using deep learning, expanding current knowledge beyond human capabilities.

8. AI offers the advantage of finding insights and expanding knowledge beyond human capabilities.
9. AI enables researchers to uncover meaningful insights from the massive amount of data generated in clinical trials, surpassing human capabilities.
10. Knowledge silos and experiential information can make it difficult to find and document important insights manually.
11. The platform connects employees with subject matter experts, leading to faster problem-solving, better decision-making, and successful clinical trials.
12. AI can automate tasks in clinical trials, reducing human involvement in creating analytics and documenting large volumes of trial data.
13. AI-driven study design optimizes trial parameters, reducing amendments and enhancing efficiency.
14. AI algorithms can identify suitable trial participants based on their medical characteristics, potentially expanding the pool of eligible participants.
15. Continuous monitoring of patients using AI ensures safety and enables timely intervention if issues arise.
16. AI tools like smartphone alerts and chatbots improve patient adherence and retention in clinical trials.
17. AI automation accelerates the summarization and analysis of large volumes of information required for regulatory submissions.
18. AI models predict drug pharmacokinetics, helping optimize dosing regimens in clinical trials.

- **Animal models:**

Animal models in drug development often fail to replicate human physiology, leading to high rates of failure in clinical trials. Relying less on animal

models and utilizing artificial intelligence (AI) can improve animal welfare, speed up drug development, and reduce costs. AI and machine learning can predict a drug's efficacy, safety, and uptake in preclinical studies, helping researchers make informed decisions and optimize testing strategies. By using AI, researchers can minimize the number of failures early in the drug development process, saving time and resources. Implementing AI in drug development can improve drug development for both animals and humans, leading to more effective and safer medication. Incorporating AI and ML into preclinical drug development allows for optimization of formulation and dosage by gathering data on a drug's efficacy and safety. AI-ML can be used to describe and diagnose trends in data, identify outliers, and assess quality in preclinical drug development. The true value of AI-ML in preclinical drug development lies in its ability to predict outcomes of experiments and inform scientists on which experiments to conduct and conditions to study. AI-ML has the potential to revolutionize preclinical drug development by complementing traditional strategies and enabling faster and more accurate prediction of drug safety and optimal formulation. Animal models have been historically important in preclinical drug development, but there are now alternatives that can accurately replicate human anatomy and physiology. Mathieu Vinken, a professor in vitro and mechanistic toxicology, advocates for combining innovative in vitro methods with computer modeling to assess drug safety. Many aspects of preclinical research can be performed without animal testing, or even improved with modern alternatives. Vinken believes that AI and machine learning are critical in extracting useful insights from existing toxicological data. The goal is to reduce the reliance on animal experiments by utilizing advanced technologies and data analysis techniques



- **Understanding disease mechanism:**

Disease modeling and target discovery are crucial initial steps in the drug discovery process and significantly impact on the success of drug development. Given the advantages of analyzing large datasets and complex biological networks, artificial intelligence (AI) is playing a growing role in modern drug target identification. We discuss the use of deep learning models for target discovery, AI-identified targets validated through experiments, and the use of synthetic data produced using generative AI for target identification. Novelty, in addition to druggability and toxicity, is a crucial factor in target selection. There is a trade-off between choosing high-confidence and novel targets. Over the past few years several AI-derived drugs have entered clinical trials, signaling the dawn of a new era in AI-driven drug discovery. AI harnesses advanced computational algorithms, machine learning, and data analysis techniques to aid researchers in understanding disease causes and accelerating the discovery of treatments. AI integrates and analyzes diverse data sources to identify patterns and correlations, enabling the identification of potential therapeutic targets and insights into disease mechanisms. AI helps prioritize potential therapeutic targets based on relevance, druggability, and likelihood of success, by analyzing data on target expression, function, and interactions. AI techniques analyze scientific literature and databases to extract valuable knowledge about disease mechanisms and potential target-disease associations, aiding in knowledge discovery and hypothesis generation. AI constructs and analyzes complex biological networks to understand the interconnectedness of molecules and pathways involved in disease, identifying key nodes and interactions for therapeutic interventions. Advances in computational technology have allowed for increased exploration of the chemical space. AI has shown potential to enhance drug

discovery methods, reducing time and cost, and increasing success rates. It can assist in identifying drug targets, screening compounds, and accelerating the design of appropriate drugs. AI algorithms can significantly reduce the time and cost to bring a drug to market. The use of artificial intelligence (AI) in medicinal chemistry has the potential to revolutionize the pharmaceutical industry by accelerating and improving the drug discovery process. AI techniques such as machine learning (ML) and natural language processing can analyze large amounts of data more efficiently and accurately, leading to more effective drug compound predictions. Deep learning (DL) has been successfully used to predict the efficacy of drug compounds with high accuracy, showcasing the potential of AI in drug discovery. AI-based methods have also shown promise in predicting the toxicity of drug candidates, further enhancing the efficiency and effectiveness of the drug discovery process. Despite the advantages of AI in developing new bioactive compounds, ethical considerations and further research are necessary to fully understand its limitations and ensure responsible use.

- **Commercialization**

EMA/FDA review and approval:

The European Medicines Agency (EMA) has released a draft reflection paper on the use of artificial intelligence (AI) in the development, regulation, and use of medicines. The paper emphasizes the importance of a human-centric approach and compliance with existing legal requirements, ethics, and fundamental rights in the use of AI and machine learning (ML) throughout the medicine lifecycle. The document is part of the joint efforts of the HMA-EMA Big Data Steering Group to enhance the European Medicines Regulatory Network's capabilities in data-driven regulation. AI and ML tools have the potential to support data acquisition, transformation, analysis, and interpretation across the medicinal product



lifecycle, including clinical trials and post-authorization activities. The European Medicines Agency (EMA) includes the use of AI in its Regulatory Science Strategy to 2025, with a focus on building regulatory frameworks and guidelines on AI validation and assessment in collaboration with academia and expert centers. The Italian Regulatory Agency (AIFA) has identified the safety of clinical trial participants and the validity of clinical data as major risks of AI/ML incorporation. AIFA requires a benefit-risk assessment to be submitted, demonstrating the usefulness and value of AI integration, as well as the reliability of the algorithm for evidence generation. The US FDA's Digital Health Center of Excellence (DHCoE) has released guidance on the development and certification of Software as a Medical Device (SaMD), applicable to AI/ML solutions used. The FDA has released a discussion paper outlining the transformative potential of AI in drug development, covering applications in target identification, compound screening, clinical research, post-market safety surveillance, and pharmaceutical manufacturing. AI and computational models are increasingly utilized in drug development, offering a cost-effective approach through Computer-aided Drug Design (CADD) techniques such as computational chemistry and molecular modeling. The FDA recognizes the increased use of AI/ML throughout the drug development life cycle and its potential to accelerate the development of safe and effective drugs, with over 100 submissions referencing AI/ML in 2021. The FDA has established the CDER AI Steering Committee (AISC) to coordinate efforts around AI/ML uses in therapeutic development and is developing. Additionally, CDER has implemented the IStand Pilot Program to expand the types of drug development tools (DDTs) included in qualification programs, including those that utilize AI/ML technology. AI/ML applications can be

used to interpret and analyze traditional DDTs, such as biomarkers and clinical outcome assessments, potentially accelerating the development of new therapeutics. FDA's CDER and CBER have established the MIDD Pilot Program to facilitate the development and application of exposure-based, biological, and statistical models derived from nonclinical and clinical data. The FDA's Sentinel Initiative, including CDER's Sentinel System, CBER's BEST system, and CDRH's NEST, are exploring AI/ML approaches to enhance postmarket safety surveillance.

Manufacturing:

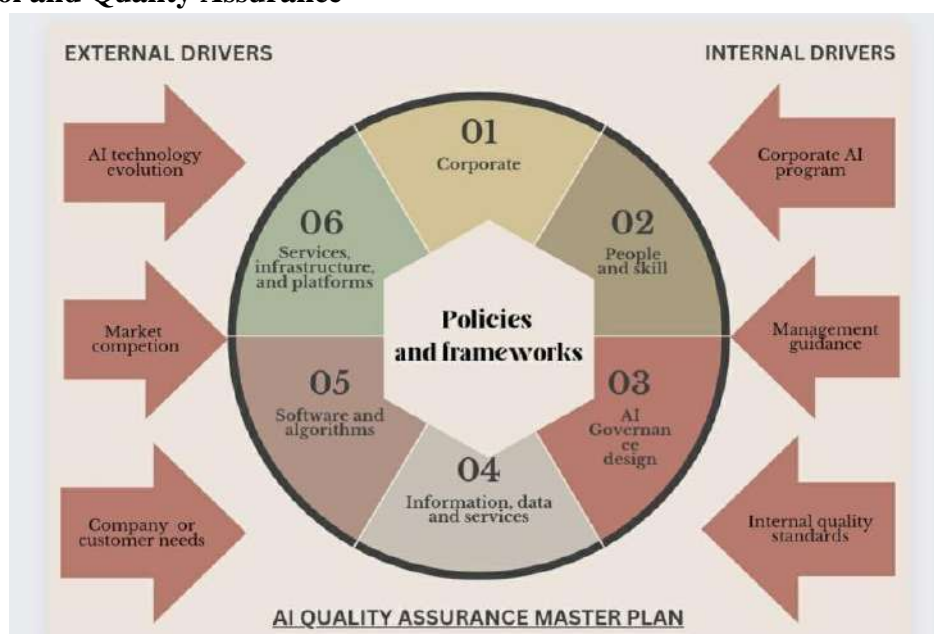
Advanced pharmaceutical manufacturing involves using methods, facilities, and controls to ensure drugs meet safety and effectiveness requirements. AI/ML can be used in pharmaceutical manufacturing for risk-based supply chain modeling, business forecasting, process optimization, and complaints reduction. AI/ML analytics can enhance process control, increase equipment reliability, detect recurring problems, and prevent batch losses. AI/ML supports Industry 4.0, enabling a well-controlled, digitized ecosystem and pharmaceutical value chain. AI/ML can improve the reliability of the manufacturing supply chain through demand forecasting, production schedule analysis, disruption mitigation, and inventory optimization. AI/ML can optimize process design by using digital twins to analyze, predict, and optimize process performance. Advanced process control (APC) allows for dynamic control of the manufacturing process to achieve desired outputs and optimize efficiency. AI/ML techniques, such as neural networks, can be used to implement APC by using real-time process data as inputs. APC applications that combine physics and chemistry knowledge with AI/ML techniques are increasingly being adopted by pharmaceutical manufacturers. Smart monitoring and maintenance using AI/ML



methods can automate and monitor manufacturing processes in real time, reducing downtime and improving efficiency. Computer vision-based quality control can analyze images to detect deviations and ensure product quality, augmenting human visual inspection. AI/ML can be used to analyze manufacturing-related deviation trends, cluster problem areas, and prioritize areas for continual improvement. AI/ML can improve manufacturing efficiency by increasing output, reducing waste, and enabling informed decision-making and quality control. Current analysis of process deviations is time-consuming and relies on manual review by quality personnel. AI/ML can

assist in analyzing large volumes of data or text to identify deviation trends and prioritize areas for improvement. By integrating AI/ML with process performance and capability metrics, manufacturing operations can be proactively monitored for trends and out-of-control events. AI/ML can also predict thresholds for triggering evaluations of corrective and preventive actions (CAPA) effectiveness, expediting the identification of root causes. Overall, AI/ML offers the potential to enhance manufacturing effectiveness and efficiency through trend monitoring, cluster problem area identification, and proactive continual.

Quality control and Quality Assurance



Fig(4)AI Quality Assurance Master Plan

AI automates routine tasks in quality assurance, such as analyzing outputs and highlighting compliance issues, improving efficiency and effectiveness. It also assists with administrative tasks like coordinating manufacturing services and checking compliance requirements. AI improves product quality while reducing costs, as quality is subjective and there is always a cost-benefit tradeoff. AI mimics human assistants to reduce costs and increase efficiency in quality control processes. AI is crucial in drug quality control,

using computer vision algorithms to identify defects and ensure compliance with quality standards. It also analyzes sensor data for predictive maintenance and detects fraud in sales and distribution data to safeguard drug distribution channels. AI plays a crucial role in fraud detection in drug distribution channels by analyzing sales and distribution data to identify suspicious patterns. Human intervention is necessary to maintain consistency and conduct quality control testing in pharmaceutical production, highlighting

the need for AI implementation. The FDA has modified manufacturing practices to better understand the processes and standards that determine the quality of pharmaceutical products. Decision trees created using a combination of human effort and AI analysis of production data can guide manufacturing cycles. Artificial neural networks (ANN) can accurately predict batch-to-batch consistency in pharmaceutical formulations. AI can regulate in-line manufacturing processes to achieve target product standards, such as monitoring freeze-drying processes. Electronic Lab Notebooks combined with intelligent algorithms can ensure product quality through automated data input, the quality of the product, but AI can assist in streamlining and optimizing the quality assurance processes. AI can analyze large amounts of data from various sources, such as lab tests, production records, and customer feedback, to identify trends and patterns that may indicate quality issues. This allows for early detection and prevention of potential problems, ensuring that only high-quality products reach the market. AI can also help in improving the efficiency of quality control processes by automating repetitive tasks, such as data entry and analysis. This frees up human resources to focus on more complex and critical tasks, ultimately improving overall productivity and effectiveness. Furthermore, AI can assist in ensuring regulatory compliance by continuously monitoring and analyzing data to identify any deviations from quality standards. Coupling A.I. and Q.C. allows for more efficient analysis of process data, leading to improved decision-making and problem-solving in the pharmaceutical industry. A.I. and Q.C. can rationalize the design of experiments, helping researchers identify the most effective approaches for drug discovery and development. By combining A.I. and Q.C., scientists can discover new molecular targets and materials,

opening up possibilities for innovative treatments and therapies.

Post release monitoring

AI can improve the efficiency of pharmacovigilance in post-market drug monitoring, ensuring compliance and accelerating customized drug guidance. GPVAI, a company, is currently using AI and OCR technology to analyze and monitor worldwide drug side effects reports. AI algorithms can analyze large datasets of patient data to detect rare or unexpected side effects that may not have been identified during clinical trials. Real-time monitoring using AI can identify potential safety issues as they arise, allowing for faster interventions. AI algorithms can predict the risk of adverse events based on patient characteristics, helping to identify high-risk patients and target interventions. AI can identify patterns and relationships between different drugs, predicting which combinations may be more likely to cause side effects. AI techniques analyze formulation parameters, drug properties, and system characteristics to model and predict drug release profiles, optimizing drug delivery systems for desired outcomes. By analyzing patient-specific data like genetics and medical history, AI can personalize drug release, tailoring delivery systems to match individual needs and enhance treatment efficacy while minimizing side effects. Machine learning algorithms, such as neural networks and support vector machines, excel at recognizing intricate patterns within datasets, enabling accurate predictions of drug release kinetics. AI models integrate diverse data sources, including physicochemical properties of drugs, formulation characteristics, and in vitro and in vivo data, providing a comprehensive understanding of drug release dynamics. Incorporating AI into drug formulation streamlines the optimization process, predicting release patterns and reducing the number of Post approval, While AI- created drugs have not gained



FDA approval to enter the market, encouraging developments have materialized within the clinical trial landscape. As of March 2022, Boston Consulting Group reported that biotechnology firms using AI as a primary strategy had over 150 small-molecule drugs in the discovery phase, with over 15 already advancing through clinical trials. Below, we list two prominent drugs that were FDA-approved for clinical trials. 1) DSP-1181, the first AI- designed drug to enter clinical trials, was developed by Exscientia and Sumitomo Dainippon Pharma. It aimed to treat obsessive-compulsive disorder and reached clinical testing in just 12 months; a fraction of the time traditional methods takes. Despite its promising start, DSP-1181 did not progress past Phase I of clinical trials and was discontinued in July 2022. It did not meet the evaluation criteria during its Phase I study, highlighting the challenges and uncertainties of AI- created drugs in the early stages of development. 2) Insilco Medicine, a biotech company in Hong Kong, has developed the world's first AI-designed anti-fibrotic drug, INS018-055, which is currently being tested in human patients. Unlike other AI-driven medications, INS018-055 was both discovered and designed using AI technology. The drug has completed initial testing and early phases, and Phase II clinical trials began in June 2023. The trials are being conducted in the U.S. and China to evaluate the safety, tolerability, pharmacokinetics, and efficacy of INS018-055 in patients with idiopathic pulmonary fibrosis. Despite promising advancements, there are currently no FDA-approved AI-created drugs, and it remains to be seen if INS. Many AI-designed drugs are still in the early stages of development or clinical testing due to the FDA's requirement for a significant amount of data from clinical trials before approval. Transparency issues may arise with AI-designed drugs as the complexity of the AI algorithm used may hinder a clear understanding of the drug's mechanism of action,

making it challenging to predict its safety and efficacy in humans. Stringent regulatory requirements imposed by the FDA could slow down the approval process for AI-designed drugs, as they may require new manufacturing processes or quality control measures that need validation before reaching the market. FDA is dedicated to ensuring the safety and effectiveness of drugs while promoting innovation in their development, including the use of AI/ML. The use of AI/ML in drug development presents both opportunities and challenges, and FDA is working to address these challenges through an agile regulatory ecosystem. FDA's CDER, CBER, and CDRH have issued a discussion paper to engage stakeholders and explore considerations for the use of AI/ML in drug and biological product development. The agency is actively seeking feedback from stakeholders as it continues to advance regulatory science in the use of AI/ML. FDA recognizes the critical role that AI/ML will play in drug development and plans to establish a flexible risk-based regulatory framework that promotes innovation while ensuring patient.

Reinforcement learning

Deep generative neural networks are increasingly used in computational chemistry for designing molecules with desired properties. Many deep learning approaches in this field use reinforcement learning to optimize the target properties of the generated molecules. However, the problem of sparse rewards often hampers the success of this approach, as most generated molecules are predicted as inactive. To address this issue, several technical innovations are proposed to improve the balance between exploration and exploitation modes in reinforcement learning. In a proof-of-concept study, a deep generative recurrent neural network architecture enhanced by these technical tricks is used to design inhibitors of the epidermal growth factor (EGFR). The potency of these inhibitors is further experimentally validated. The



proposed technical solutions are expected. Deep-generative models are a crucial focus in Computer-Assisted Drug Discovery (CADD) for designing molecules with desired properties. These models use different molecular representations, such as SMILES strings and molecular graphs, to generate novel molecules. Several models have been proposed for generating SMILES strings and molecular graphs of synthetically feasible molecules. Initially, these models are trained on diverse datasets to generate a wide range of molecules. A naïve generative model refers to a model trained on a generic dataset before any property optimization. Olivecrona et al. and Blaschke et al. proposed reinforcement learning algorithms to maximize the predicted activity of being used in computational chemistry to design molecules with desired properties. Many deep learning approaches use reinforcement learning to optimize the properties of the generated molecules. Sparse rewards are a challenge in this approach, as most generated molecules are predicted to be inactive. Several technical innovations are proposed to improve the balance between exploration and exploitation in reinforcement learning. A proof-of-concept study demonstrates the application of these technical tricks to design inhibitors of the epidermal growth factor (EGFR). The potency of these inhibitors is experimentally validated. The proposed technical solutions are expected to improve the success rate of finding novel bioactive compounds using generative and reinforcement learning. Deep-generative models are being used in Computer-Assisted Drug Discovery (CADD) to design molecules with desired properties. These models

Machine learning

generated molecules against specific receptors, while Jin et al. used RationaleRL to maximize activity against specific inhibitors. Born et al. suggested using reinforcement learning on merged protein/ligand latent space constructed by the VAE, but no experimental validation was conducted in these studies. Zhavoronkov et al. not only proposed a novel generative tensorial reinforcement-learning algorithm but also validated their method by designing potent DDR1 kinase inhibitors. Most theoretical studies on de novo molecular design focus on optimization tasks for properties like LogP and QED, using scoring functions to assign rewards to generated molecules. Deep generative neural networks are

use different molecular representations, such as SMILES strings and molecular graphs, to generate novel molecules. Initially, the models are trained on a diverse dataset to generate a broad distribution of molecules. Reinforcement learning (RL) is used to optimize the properties of the generated molecules. Examples of RL algorithms include REINVENT, memory-assisted RL, and Rationale RL. Some studies have used RL to maximize the predicted activity of molecules against specific receptors or enzymes. Experimental validation of the computational hits is important, as shown by Zhavoronkov et al. who designed potent kinase inhibitors and performed. Theoretical studies on de novo molecular design often use optimization tasks for properties like Log P21 and Quantitative Estimate of Drug likeness (QED) 22. These tasks use objective metrics obtained from a molecule's.



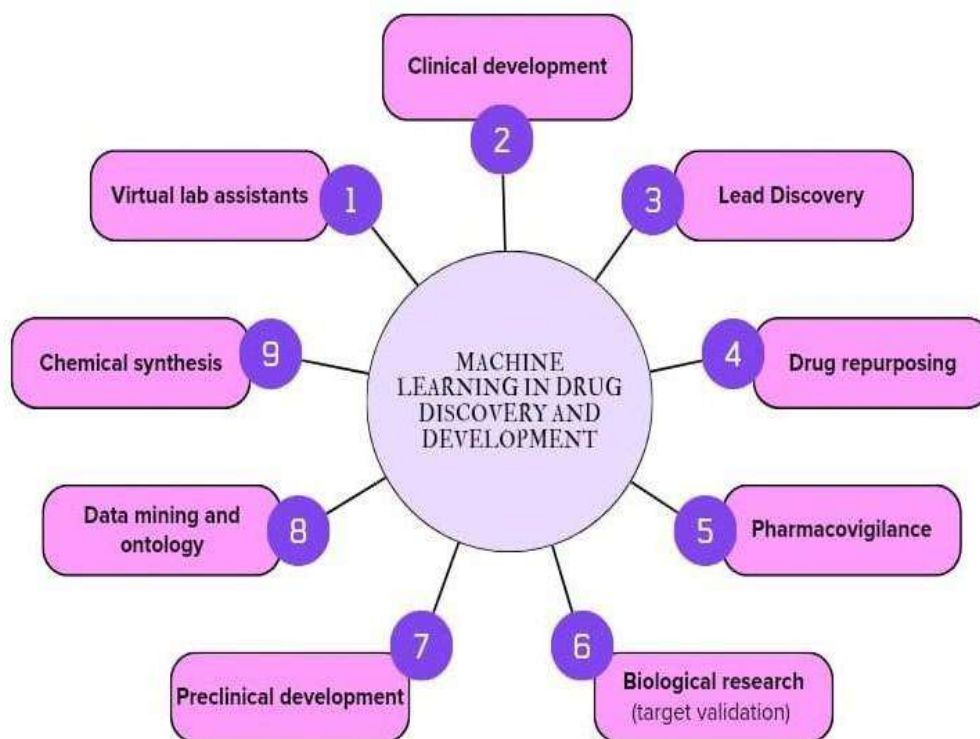


Fig.no. 5 Machine learning drug discovery and development

The ML process is similar to the learning behavior of the human brain, interpreting various perceptions through billions of neurons. Machines analyze and calculate data using an electronic nose, similar to how the human brain interprets images, sound, smell, structures, movements, and dimensional patterns. The ML process requires two components: input (data for analysis) and output (result of ML algorithm calculation). High-quality data is prepared for ML by randomizing and excluding anomalies or duplicates. The data is divided into training, validation, and test sections. Before testing different datasets, training and validation are performed, followed by algorithm, and learning model selection. The selection of algorithm and learning model depends on the type of data and the task to be automated. ML is divided into three models: supervised learning, unsupervised learning, and reinforcement learning, based on the type of data. In supervised learning, well-labelled input data is required for learning, and the algorithm can analyze non-labelled data after training. In this model, elements

are sorted into groups with pre-defined features, and the value is predicted based on training data calculations. Unsupervised learning focuses on finding patterns and correlations between non-labelled data points. An algorithm must assemble data by characteristics that differentiate them from other groups of objects. Data types like MRI scans, digital photographs, and audio signals have high dimensions indicating the number of features for each observation. ML reduces the amount of data by selecting important attributes or combining.

AI based software tools for drug development process

AI encompasses various method domains, including reasoning, knowledge representation, solution search, and machine learning (ML). ML utilizes algorithms to identify patterns in classified data, with deep learning (DL) being a subfield that employs artificial neural networks (ANNs). ANNs are interconnected computing elements, resembling human neurons, and simulate the transmission of electrical impulses in the brain. ANNs consist of nodes that receive separate inputs

and convert them to output, using algorithms to solve problems. Different types of ANNs include multilayer perceptron (MLP) networks, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), which can be trained using supervised or unsupervised methods. MLP networks are used for pattern recognition, optimization aids, process identification, and controls. They are trained through supervised training procedures and can function as universal pattern classifiers. RNNs are networks with a closed loop that can memorize and store information. They are used for tasks such as Boltzmann constants and Hopfield networks. CNNs are dynamic systems with local connections, commonly used in image and video processing, biological system modeling, and pattern recognition. Other complex forms of AI networks include Kohonen networks, RBF networks, LVQ networks, counter-propagation networks, and ADALINE networks. IBM Watson supercomputer, developed using AI technology, is a powerful tool for analyzing medical information and suggesting treatment strategies for cancer patients. The system can also rapidly detect diseases, as demonstrated by its ability to detect breast cancer in just 60 seconds. BM Watson's vast database allows it to correlate patient data with relevant information, providing valuable insights for personalized treatment plans. The use of AI technology in healthcare, such as IBM Watson, has the potential to revolutionize the way diseases are diagnosed and treated. The rapid detection capabilities of IBM Watson can significantly improve patient outcomes by enabling early intervention and timely treatment.

Table.no. 1 Examples of AI tools used in drug discovery

Table	Details
Deep Chem	MLP model that uses a python-based AI system to find a suitable candidate in drug discovery.

Deep Tox	Software that predicts the toxicity of total of 12000 drugs.
Deep Neural NetQSAR	Python-based system driven by computational tools they aid detection of the molecular activity of compounds.
ORGANIC	A molecular generation tool that helps to create molecules with desired properties.
PotentialNet	Uses NNs to predic binding affinity of ligand.
Hit Dexter	ML technique to predict molecular that might respond to biochemical assays.
DeltaVina	As scoring function for rescoring Drug-ligand binding affinity.

Model architecture

Traditional machine learning models were commonly used before the rise of deep learning in tasks such as predicting drug likeliness, physicochemical properties, pharmacokinetic parameters, and pharmacodynamics properties. Score-based classification models like support vector machines (SVM) and K nearest neighbors (KNN) were used, as well as probability-based classification models like random forest (RF), naive bayes (NB), and logistic regression (LR). Despite the success of traditional machine learning models, deep neural networks (DNNs) have shown superior performance in various tasks. Big data application requires collaboration between "Big data analytics" and "Implementation of data modeling" processes using Model-driven architecture (MDA). Data modeling in MDA should be done quickly to verify the data model and discover new data resources for services. The research goal is to predict the side effects of drugs using data mining methods at the intersection of statistics, machine learning, and database systems. Data modeling can also lead to the discovery of new uses for old drugs in the field of drug discovery. A prototype system has been developed to implement the prediction model and data model for drug discovery and test their practicality.



ADME studies are crucial in drug discovery, providing important information on the pharmacokinetic and pharmacodynamic properties of potential drug candidates. The mouse-liver microsomal (MLM) assay is a key pharmacological assay that predicts metabolic stability and is required for any investigational new drug (IND) application. Computational tools/models that can predict the results of ADME assays, including MLM, would greatly accelerate the drug discovery process and reduce costs. Dr. Yufeng Tseng and his team from National Taiwan University compared machine learning models predicting MLM stability to a model constructed using deep learning methods in a recent review. The review highlights the potential of deep learning methods in accurately predicting MLM stability, offering a promising approach for improving drug deep learning models are a subset of machine learning methods that use artificial neural networks inspired by biological neural networks. Neural networks consist of artificial neurons connected to each other through edges, allowing them to transmit information. Nodes in neural networks are organized into layers, including input nodes, hidden layers, and an output layer. The learning process for neural networks involves adjusting the weights and biases of nodes to correctly classify features within the input layer. Deep learning models offer advantages over traditional machine learning models, such as faster training times, efficient computation, and higher predictive accuracy on diverse datasets. Graph convolutional neural networks (GCNs) have made deep learning models more practical in computational chemistry by directly inputting graphical features of molecules into the neural network. Unlike traditional molecule representation methods that use fixed-dimensional feature vectors, GCNs can develop their own flexible and machine-optimized fingerprints, leading to stronger predictive performance and

more streamlined feature representation. In a study on MLM stability prediction, a GCN outperformed a state-of-the-art Bayesian classifier in terms of accuracy, sensitivity, specificity, and AUC ROC scores. The GCN showed increased prediction accuracy, sensitivity, specificity, and AUC ROC scores compared to other models, with a 6.6% increase in overall accuracy in the test set and a 22.6% increase in the validation set. The GCN also had higher ROC scores, with a +0.076 increase in the validation set and a +0.160 increase in the test set. The C5.0 decision tree model, trained on the same dataset, showed overfitting and decreased accuracy compared to the GCN. Other machine learning models, including Bayesian classifiers and random forest models, also had decreased accuracy compared to the GCN. The C5.0 classification model developed by the researchers showed overfitting.

CONCLUSION

AI tools are being used in the pharmaceutical industry to improve drug development, personalized medications, and resource utilization. While concerns about job losses and regulations exist, AI can assist in quick decision-making and create new chemical entities. The revenue from AI-based solutions in the pharmaceutical sector is projected to reach \$2.9 billion by 2024. The synergy between human minds and AI is crucial for successful drug discovery.

REFERENCE

1. Askin, S., Burkhalter, D., Calado, G. and El Dakrouni, S., 2023. Artificial Intelligence Applied to clinical trials: opportunities and challenges. *Health and Technology*, 13(2), pp.203-213.
2. Harrer, S., Shah, P., Antony, B. and Hu, J., 2019. Artificial intelligence for clinical trial design. *Trends in pharmacological sciences*, 40(8), pp.577-591.
3. Harrer, S., Shah, P., Antony, B. and Hu, J., 2019. Artificial intelligence for clinical trial



- design. *Trends in pharmacological sciences*, 40(8), pp.577-591.
- Naik, N.N., Vadloori, B., Poosala, S., Srivastava, P., Coecke, S., Smith, A., Akhtar, A., Roper, C., Radhakrishnan, S., Bhyravbhatla, B. and Damle, M., 2023. *Advances in Animal Models and Cutting-Edge Research in Alternatives: Proceedings of the Third International Conference on 3Rs Research and Progress*, Vishakhapatnam, 2022. *Alternatives to Laboratory Animals*, p.02611929231180428.
 - Balakrishnan, A., Jaglan, P., Selly, S., Kumar, V. and Jabalia, N., 2023. Emerging trends of blockchain in bioinformatics: A revolution in health care. *Distributed Computing to Blockchain*, pp.389-404.
 - Pun, F. W., Ozerov, I. V., & Zhavoronkov, A. (2023). AI-powered therapeutic target discovery. *Trends in pharmacological sciences*, 44(9), 561–572. <https://doi.org/10.1016/j.tips.2023.06.010>
 - Bhattamisra, S.K., Banerjee, P., Gupta, P., Mayuren, J., Patra, S. and Candasamy, M., 2023. Artificial Intelligence in Pharmaceutical and Healthcare Research. *Big Data and Cognitive Computing*, 7(1), p.10.
 - Monson, T.A., 2020. Risk attitudes within 'complex youth' assessment and decision making: Professional perspectives. *Childcare in Practice*, 26(2), pp.210- 222
 - Kumar, M., Nguyen, T.N., Kaur, J., Singh, T.G., Soni, D., Singh, R. and Kumar, P., 2023. Opportunities and challenges in application of artificial intelligence in pharmacology. *Pharmacological Reports*, 75(1), pp.3-18.
 - Karpatne, A., Kannan, R. and Kumar, V. eds., 2022. *Knowledge Guided Machine Learning: Accelerating Discovery Using Scientific Knowledge and Data*. CRC Press.
 - Son, W.S., 2018. Drug discovery enhanced by artificial intelligence. *Biomedical Journal of Scientific & Technical Research*, 12(1), pp.8936-8938.
 - Vidhya, K.S., Sultana, A., Kumar, N., Rangareddy, H., Vidhya, K.S. and Madalageri, N.K., 2023. Artificial Intelligence's Impact on Drug Discovery and Development from Bench to Bedside. *Cureus*, 15(10).
 - Korshunova, M., Huang, N., Capuzzi, S., Radchenko, D.S., Savych, O., Moroz, Y.S., Wells, C.I., Willson, T.M., Tropsha, A. and Isayev, O., 2022. Generative and reinforcement learning approaches for the automated de novo design of bioactive compounds. *Communications Chemistry*, 5(1), p.129.
 - Pun, F.W., Leung, G.H.D., Leung, H.W., Rice, J., Schmauck-Medina, T., Lautrup, S., Long, X., Liu, B.H.M., Wong, C.W., Ozerov, I.V. and Aliper, A., 2023. A comprehensive AI-driven analysis of large-scale omic datasets reveals novel dual-purpose targets for the treatment of cancer and aging. *Aging Cell*, p.e14017.
 - Bleicher, L.S., van Daelen, T., Honeycutt, J.D., Hassan, M., Chandrasekhar, J., Shirley, W., Tsui, V. and Schmitz, U., 2022. Enhanced utility of AI/ML methods during lead optimization by inclusion of 3D ligand information. *Frontiers in Drug Discovery*, 2, p.1074797.

HOW TO CITE: Somnath S. Davkhar, Devadhe Sanika S., Khedkar Namdev Bhagwan, A Review Of Artificial Intelligence In Drug Discovery And Development, *Int. J. of Pharm. Sci.*, 2024, Vol 2, Issue 3, 7-22. <https://doi.org/10.5281/zenodo.10774652>

