Global Journal of Medical Research: B
Pharma, Drug Discovery, Toxicology & Medicine
Volume 25 Issue 1 Version 1.0 Year 2025
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4618 & Print ISSN: 0975-5888

Drug Development and Discovery Considering Artificial Intelligence: A Through Analysis

By Sachin Verma, Ridam Gaud, Ankita Bhattacharjee & Sadia Parveen

Sharda University

Abstract- Artificial Intelligence (AI) is increasingly reshaping drug discovery and development by offering new computational capabilities that significantly enhance efficiency, accuracy, and innovation. This comprehensive review discusses the evolving role of AI across various stages of pharmaceutical R&D—from early target identification and validation to lead optimization, preclinical assessment, and clinical trials. With the growing complexity and costs associated with traditional drug development pipelines, AI presents powerful alternatives through machine learning (ML), deep learning (DL), and natural language processing (NLP) tools that enable rapid data analysis, compound generation, and predictive modeling.

In target discovery, AI algorithms analyze vast omics datasets to identify novel biological targets, while virtual screening models streamline high-throughput screening of chemical libraries with improved hit rates.

Keywords: *clinical trials, virtual screening, deep learning, machine learning, artificial intelligence, drug discovery, target identification, pharmaceutical innovation.*

GJMR-B Classification: *NLMC: R855.3, Q335, RS420*

Strictly as per the compliance and regulations of:

Drug Development and Discovery Considering Artificial Intelligence: A Through Analysis

Sachin Verma [α], Ridam Gaud [σ], Ankita Bhattacharjee [ρ] & Sadia Parveen [ω]

Abstract- Artificial Intelligence (AI) is increasingly reshaping drug discovery and development by offering new computational capabilities that significantly enhance efficiency, accuracy, and innovation. This comprehensive review discusses the evolving role of AI across various stages of pharmaceutical R&D—from early target identification and validation to lead optimization, preclinical assessment, and clinical trials. With the growing complexity and costs associated with traditional drug development pipelines, AI presents powerful alternatives through machine learning (ML), deep learning (DL), and natural language processing (NLP) tools that enable rapid data analysis, compound generation, and predictive modeling.

In target discovery, AI algorithms analyze vast omics datasets to identify novel biological targets, while virtual screening models streamline high-throughput screening of chemical libraries with improved hit rates. Lead optimization benefits from AI's ability to predict ADMET (absorption, distribution, metabolism, excretion, and toxicity) profiles, thus reducing the failure rate in later stages. In clinical research, AI assists in patient stratification, real-time monitoring, and biomarker identification, accelerating trial timelines and enhancing patient safety.

This review evaluates prominent AI platforms such as DeepChem, AtomNet, Schrödinger's suite, and AlphaFold, along with case studies from industry leaders like Pfizer, Novartis, and Insilco Medicine. Challenges such as data quality, model interpretability, algorithmic bias, and regulatory concerns are critically analyzed. The paper concludes by identifying future research opportunities and emphasizing the need for collaborative frameworks between AI developers, biologists, and regulatory bodies.

As AI continues to evolve, its integration into the drug discovery lifecycle holds the promise of significantly transforming pharmaceutical innovation, enabling more targeted therapies and advancing the vision of precision medicine.

Keywords: clinical trials, virtual screening, deep learning, machine learning, artificial intelligence, drug discovery, target identification, pharmaceutical innovation.

I. Introduction

Drug development and discovery is a time-consuming, costly, and intricate process that often takes decades and billions of dollars. The average cost of introducing a novel medication to market, including opportunity and direct expenses, is about $2.6 billion, according to the Tufts Centre for the Study of medicine Development [1]. Despite these substantial investments, the overall success rate of drug candidates from Phase I trials to market approval is below 10% [2]. This high attrition rate in later phases, combined with escalating research and development (R&D) expenditures, has driven the pharmaceutical industry to seek more efficient and accurate methods.

Artificial intelligence (AI) has emerged as a disruptive force in this field thanks to its powerful computational capabilities. The methods of computing that are included in the artificial intelligence (AI) category include machine learning (ML), deep learning (DL), neural networks, reinforcement learning, and natural language processing (NLP). These technologies can process large biomedical datasets and find patterns that conventional techniques are unable to detect[3,4]. Many facets of drug development, including target identification, chemical creation, protein structure prediction, and biomarker analysis, are currently being addressed by these methods.

For instance, AlphaFold by DeepMind has demonstrated the potential of AI by accurately predicting protein 3D structures, which are critical in understanding biological functions and designing novel drugs [5]. Additionally, the integration of AI with screening at high throughput and omics technologies has enabled pharmaceutical companies to significantly shorten discovery cycles and reduce cost burdens.

This review aims to comprehensively examine the ways that AI is transforming the pharmaceutical industry, focusing on applications across all phases of the development and discovery of new drugs. We present an overview of the AI-powered drug discovery pipeline, analyze current tools and platforms, and explore into the main technical, ethical, and regulatory issues. We also provide insight into the future direction of AI in pharmaceutical R&D, highlighting its significance inprecision medicine and collaborative innovation.

Optimising Clinical Trials with AI: Artificial intelligence is quickly changing how clinical trials are planned and carried out by enhancing patient recruitment, stratification, and real-time monitoring. Natural Language Processing (NLP) algorithms are being used to mine clinical notes and electronic health records (EHRs) in order to find qualified participants with high

Author α: Assistant Professor, Department of Pharmaceutics, Sharda School of Pharmacy, Sharda University, Agra, UP.
Author σ: UG Student, Sharda School of Pharmacy, Sharda University, Agra, UP.
Author ρ: UG Student, Siddhartha School of Pharmacy, Dehradun, Uttarakhand.
Author ω: Assistant Professor, Department of Pharmaceutics, Jaipur National University, Jaipur, Rajasthan. e-mail: Sadia.p664@gmail.com

precision, thereby addressing a significant bottleneck in clinical research: patient recruitment [55]. Additionally, AI makes predictive modelling for patient stratification possible, which enables researchers to divide up patients according to how they are likely to respond to treatment. This reduces trial size and increases statistical power [56]. For instance, IBM Watson Health and Deep6 AI have shown that they can use AI-based analytics to cut patient screening times by over 80% [57]. AI-enabled wearable technology also makes it easier to monitor trial participants in real time, guaranteeing early adverse event detection and enhancing compliance by identifying behavioural patterns [58]. These applications make clinical trials more flexible and effective by lowering trial costs and schedules while simultaneously enhancing safety and data quality.

Pharmacovigilance and AI: In order to ensure long-term drug safety, post-marketing surveillance, also known as pharma covigilance, is essential. Artificial intelligence (AI) provides significant benefits in identifying adverse drug reactions (ADRs) from disparate data sources. Pharma covigilance now uses AI algorithms that can evaluate unstructured data from social media, electronic health records, and scientific literature to discover early warning signals, replacing its previous reliance on manual signal recognition via spontaneous reporting methods [48,59]. Traditional statistical techniques such as disproportionality analysis are not as sensitive or specific in identifying possible ADRs as machine learning classifiers and natural language processing tools [33,60]. Tools such as Med Watcher Social and FAERSmine, for example, have used AI to more quickly and thoroughly detect concealed safety signs [61]. In the era of expedited medication approvals and international pharmacovigilance regulations, artificial intelligence (AI) greatly speeds up regulatory response and improves patient safety by automating signal detection and prioritisation.

AI for Repurposing Drugs: Due to AI-driven approaches, the process of developing new beneficial applications for licensed drugs—known as drug repurposing—has quickened. Through the integration of omics data, clinical phenotypes, and chemical structural information, artificial intelligence systems have the ability to reveal previously undiscovered relationships between medications and illnesses [62]. To make highly accurate predictions about drug-disease correlations and biological pathways, platforms such as Benevolent AI and GENTRL use deep learning models [63]. AI-powered repurposing was particularly important during the COVID-19 pandemic, when medications such as baricitinib were discovered as possible treatments using machine learning algorithms just a few weeks after the virus was characterised [64]. These techniques shorten the discovery timeline and reduce the risk of failure by focusing on compounds with known safety profiles. As a result, AI has become a strategic tool for extending the therapeutic utility of existing drug libraries in a cost-effective and timely manner.

II. AI Applications Across the Drug Development Pipeline

a) *Target Identification and Validation*

The drug development process begins with identifying and validating biological targets associated with disease mechanisms. Systems powered by artificial intelligence (AI) are demonstrating impressive abilities in examining enormous biological data repositories to find possible therapeutic targets with previously unheard-of speed and precision. By integrating and analyzing genomic, proteomic, and clinical data, AI tools can pinpoint molecules or biological pathways that play key roles in disease progression, providing valuable insights into potential therapeutic interventions.

Machine learning approaches can predict protein structures and functions, analyze protein-protein interactions, and identify disease-relevant biological pathways. These capabilities enable researchers to discover innovative targets and repurpose existing drugs for new indications. Additionally, AI algorithms can evaluate target druggability and prioritize candidates based on their potential clinical relevance, significantly streamlining the the initial phases of discovering new drugs.

b) *Drug Design and Screening*

AI streamlines the entire procedure of determining compounds that can efficiently modify prospective targets once they have been discovered. In conventional high-throughput screening, hundreds of chemicals are tested experimentally, which takes a lot of time and resources. Large libraries of compounds can be quickly scanned using AI-guided screening approaches to find the ones that have the best chance of binding to the target.

Beyond screening existing compounds, AI enables De novo drug design, such as creating novel chemical compounds structures optimized for specific targets and properties. Generative models can create new chemical entities with desired characteristics while maintaining structural validity and synthetic feasibility. Additionally, AI algorithms aids researchers in prioritising the most promising prospects for additional development by forecasting the pharmacokinetic and pharmacodynamic characteristics of drugs.

c) *Preclinical Research*

Comprehensive safety and efficacy evaluations of possible medication candidates are part of preclinical research. In the past, this procedure has been costly, difficult, and frequently ineffective. This stage is being revolutionised by AI tools, which simplify data analysis,

anticipate drug interactions, and more effectively identify interesting molecules.

Potential drug candidates' toxicity can now be predicted using machine learning algorithms, which drastically reduces down the time and money needed for experimental testing. Preclinical testing time and expense can be significantly decreased by using AI-driven systems, enabling researchers to quickly screen thousands of potential compounds and rank those that seem to have the greatest potential. These predictive models can evaluate drug metabolism, absorption, distribution, and potential side effects, enabling earlier elimination of problematic compounds.

d) Clinical Trials

The inefficiencies of traditional clinical trial designs frequently result in exorbitant expenses, protracted schedules, and perhaps ambiguous findings. AI can enhance the discovery of new therapies, trial results, and patient recruitment by customising treatment plans and inclusion criteria using predictive analytics.

AI-powered algorithms evaluate patient data to stratify patient populations and find pertinent biomarkers, improving trial design and raising the possibility of positive results. Additionally, prior to starting expensive and time-consuming clinical trials, researchers can test hypothetical trial situations, improve study protocols, and reduce risks using AI-driven simulations. Companies are developing patient-centric solutions using wearables and apps to gather real-world data, enabling more accurate tracking of drug safety and efficacy at the individual level.

e) Post-market Surveillance and Personalized Medicine

AI continues to provide value after drug approval through post-market surveillance and personalized medicine applications. AI algorithms can monitor real-world data to detect previously unidentified side effects, evaluate long-term safety profiles, and identify additional therapeutic indications.

The development of personalised medicine is one of the most exciting uses of AI in medication development. Through better diagnoses, individualised data collection, and clinical decision support, AI can quicken this trend. AI systems are able to retain and examine patient data, such as lifestyle characteristics, clinical histories, and genetic profiles, enabling more specialised treatment strategies. Furthermore, AI technology can find biomarkers linked to medication responses or the advancement of a disease, allowing for the creation of therapies with the highest possible efficacy and the fewest possible side effects.

III. AI Techniques in Drug Discovery

a) Machine Learning Algorithms

Advanced artificial intelligence (AI)-based methods for drug development have replaced more conventional quantitative structure-activity relationship (QSAR) modelling techniques. Among these methods are decision trees, Support vector algorithms, random forests, and linear discriminant analysis (LDA), which can speed up QSAR analysis and increase prediction accuracy. Large datasets can be explored by algorithms for machine learning to discover links and patterns that traditional analysis techniques can omit.

b) Deep Learning Models

Deep learning techniques have revolutionized drug discovery by enabling more complex pattern recognition and predictive capabilities. Recurrent neural networks deal with sequential data, whereas convolutional neural networks may analyse imaging data. Graph neural networks are particularly valuable for analyzing molecular structures and predicting compound properties. Deep learning models excel at integrating diverse data types and extracting meaningful features from complex biological datasets.

c) Generative AI

Generative AI models represent a particularly exciting advancement in drug design. Novel structures for molecules with desired features can be generated by using these computational models, enabling the creation of entirely new chemical entities tailored to specific targets. Methodologies incorporating learning via reinforcement, generative networks with adversarial properties, and variational autoencoders allow researchers to explore chemical space more efficiently and design molecules with optimal characteristics.

d) Large Language Models

Emerging large language models have recently been applied to drug discovery, offering new capabilities for analyzing scientific literature, predicting molecular interactions, and generating insights from unstructured data. These models can process vast amounts of textual information, enabling researchers to leverage knowledge contained in millions of scientific publications and databases to inform drug development decisions.

Table 1: Key AI Techniques and their Applications in Drug Discovery

AI Technique	Applications for Drug Discovery	Examples
Algorithms in machine learning	QSAR modeling, property prediction, virtual screening	Random Forest, SVM, Decision Trees
Deep Learning	Protein structure prediction, binding affinity estimation	Neural networks, CNN, RNN
AI generative	Designing de novo molecules and optimising leads	GANs, VAEs, Reinforcement Learning
Large Language Models	Literature analysis, knowledge extraction	Emerging models for scientific text processing
Computer Vision	High-throughput screening analysis, histopathology	Image recognition algorithms

Natural Language Processing	Mining scientific literature, patient data analysis	Text classification, named entity recognition
Reinforcement Learning	Molecular optimization, clinical trial design	Policy-gradient methods, Q-learning

IV. Overview of the Drug Development Pipeline

The traditional drug development pipeline is a multistep, iterative process designed to identify, develop, and bring new therapeutic agents to market. The main goal is to ensure theefficacy, safety and quality of novel pharmaceutical compounds before they reach patients. Each stage presents unique challenges and opportunities for optimization, especially with taking into accountof Artificial Intelligence (AI).

Table 2: AI Applications and Emerging Software in Drug Development Stages

Drug Development Stage	AI Techniques Used	Emerging Pharmacy Software / Tools	Recent Advancements / Key Features	Description
Target Identification & Validation	- Machine Learning (ML) - Deep Neural Networks (DNN) - NLP for literature mining	- IBM Watson Discovery -BioXpress - Open Targets Platform	- Predicting druggable targets from multi-omics data - NLP-powered biomarker extraction - Integration of CRISPR screening data	Identifying genes, proteins, or pathways associated with disease pathogenesis using techniques such as genomics, transcriptomics, and bioinformatics tools.
Hit Discovery / Screening	- Virtual Screening - Reinforcement Learning - Generative Adversarial Networks (GANs)	- Schrödinger Glide -DeepChem -Atomwise (AtomNet)	- AI-guided high-throughput screening - Ligand-based drug design using DL - Molecular fingerprint learning	Screening large chemical libraries (both physical and virtual) to find compounds that exhibit biological activity against the identified target.
Lead Optimization	- QSAR Modeling - Active Learning - Bayesian Optimization	-MOSES -DeepMol - Chemprop	- AI for ADMET prediction - Property optimization with molecular generative models - Predictive SAR mapping	Refining the chemical structure of hit compounds to enhance selectivity, potency,safety, metabolic stability. This often involves structure-activity relationship (SAR) studies.
Preclinical Development	- Supervised ML - Toxicity Prediction Models - Graph Neural Networks (GNNs)	-pkCSM -DeepTox -ADMETlab 2.0	- Predicting off-target toxicity - Physiologically-based pharmacokinetics (PBPK) modeling - GNN-based prediction of toxicity profiles	Conducting in vivo and in vitro experiments to evaluate pharmacokinetics (PK), pharmacodynamics (PD), and toxicity profile of the optimized compound. Animal models are used for predictive safety.
Clinical Trial Design & Execution	- AI-based patient stratification - Predictive Modeling - Digital Twins	- Deep 6 AI - Unlearn. AI - TriNetX	- Simulated control arms using digital twins - AI for trial recruitment and dropout prediction - Adaptive clinical trial design	Testing in human volunteers across three phases: Phase I (safety and dosing), Phase II (efficacy and side effects), and Phase III (confirmation of effectiveness and monitoring of adverse reactions).
Regulatory Submission & Market Access	- Natural Language Processing - Data Harmonization - Knowledge Graphs	- FDA's CDER AI-based systems - Regulatory NLP Tools (e.g., Linguamatics)	- Automated generation of eCTD documents - Real-world evidence integration - AI for pharmacovigilance signal detection	Submission of comprehensive data packages to agencies like the FDA or EMA to demonstrate safety, efficacy, and manufacturing quality for market authorization.

In recent years, AI has begun to redefine this traditional workflow. At the target identification stage, machine learning models analyze high-throughput omics data to uncover disease-relevant genes or pathways. For example, studies such as those by Zeng et al. (2020) have demonstrated how deep learning platforms like Deep Target can effectively predict molecular targets based on multi-omics integration.

For lead discovery, AI-driven virtual screening platforms such as AtomNet and DeepChem can evaluate billions of compounds in silico, drastically reducing the time and cost associated with physical screening (Wallach et al., 2015; Ramsundar et al., 2019). These platforms use convolutional neural networks and graph-based learning algorithms to predict molecular binding affinities and activities with impressive accuracy. The use of molecular docking algorithms accelerated by AI further helps prioritize potential hits for synthesis and biological testing.

During lead optimization, AI tools predict physicochemical properties, simulate metabolism, and propose novel analogs through generative models like GANs and reinforcement learning. For instance, Insilico Medicine developed DDR1 kinase inhibitors using a deep generative approach that reached preclinical trials in less than 18 months (Zhavoronkov et al., 2019). AI

can also help guide Structure-Based Drug Design (SBDD) and Ligand-Based Drug Design (LBDD) by learning SAR patterns and predicting molecular modifications that may improve activity profiles.

In preclinical testing, AI supports toxicity prediction using in silico methods such as DeepTox and ProTox-II, trained on toxicogenomic databases. These systems can assess hepatotoxicity, cardiotoxicity, and off-target interactions, potentially reducing the need for animal testing (Mayr et al., 2016). Additionally, systems biology modeling, powered by AI, can simulate entire organ systems, helping to predict human-relevant responses without invasive methods.

For clinical trials, AI aids in optimizing trial design, predicting dropout rates, and stratifying patient populations. IBM Watson Health, for instance, has collaborated with healthcare institutions to improve patient matching for oncology trials, reducing recruitment time by over 50% (IBM Watson Health, 2020). AI also assists in remote monitoring through wearable devices and real-time data integration. Furthermore, NLP algorithms process patient records and unstructured data to identify adverse drug reactions and eligibility criteria at scale.

Regulatory approval processes are increasingly incorporating AI-driven analyses. Regulatory bodies like the U.S. FDA have released frameworks supporting the use of machine learning in regulatory submissions, and the EMA's Big Data Task Force is developing standards for validation and traceability of AI-based outputs (FDA, 2021; EMA, 2020). These initiatives aim to ensure transparency, reproducibility, and clinical relevance of AI-supported decisions in healthcare product evaluations.

Overall, the integration of AI across the drug development pipeline enhances decision-making, reduces resource waste, and improves success rates by addressing bottlenecks that have historically hindered drug development timelines and outcomes. As more validated models and regulatory frameworks are established, AI will likely become a cornerstone of pharmaceutical R&D workflows.

a) Recent Developments and Innovations (2023-2025)

The field of AI in drug discovery has witnessed significant advancements in recent years. In 2023, research company Startus Insights identified nine key trends where companies are breaking new ground, including AI & data analytics, patient-centric trials, and assay development.

Companies like Protai in Israel are leveraging AI to build drug discovery platforms, while Italian startup Netabolics is predicting drug effects by digitizing human cells. Other innovations include gut-on-chip testing solutions developed by Latvian startup Cellbox, which replicate human organs and run experiments on chips controlled by biosensors.

Researchers can now swiftly search through huge libraries of compounds for possible therapeutic candidates thanks to the combination of AI with high-throughput screening methods, greatly speeding up the early phases of drug discovery. AI analysis in conjunction with developments in proteomics, genomics, and other omics technologies has led to a better understanding of biological systems and disease pathways.

By 2025, the combination of both generative AI and enormously language modelshave further enhanced the efficiency and effectiveness of the drug development process, enabling more precise target identification, faster lead optimization, and improved clinical trial design.

b) Challenges and Limitations

Despite remarkable progress, plenty of challenges still exist in the use of AI in drug development and development. Data quality issues represent a significant constraint, as the level of accuracy of the data determines how well AI models perform., they are trained on. Limited availability of high-quality, well-annotated datasets, particularly for rare diseases and novel targets, can restrict AI model performance.

Regulatory considerations pose another challenge, as regulatory frameworks are still adapting to AI-driven drug development approaches. Demonstrating the reliability and reproducibility of AI-generated results to regulatory authorities requires careful validation and transparency.

Ethical concerns must also be addressed, particularly regarding data privacy, bias in training datasets, and the appropriate use of AI-generated findings. Ensuring that AI applications in drug discovery adhere to ethical principles is essential for maintaining public trust and promoting responsible innovation.

Technical constraints continue to limit certain AI applications, especially for complex biological systems that are not fully understood. Interdisciplinary collaboration between computational experts and domain specialists is essential to overcome these limitations and realize the full potential of AI in pharmaceutical research.

V. AI in Target Identification and Validation

Target identification and validation represent the foundational step in drug discovery, where researchers aim to pinpoint disease-associated genes, proteins, or pathways that can be modulated to achieve therapeutic effects. Traditional approaches often rely on labour-intensive methods, including gene knockout studies, proteomics, and biochemical assays. However, these methods are limited by their scalability and complexity.

Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), is transforming this process by integrating and analyzing diverse biological datasets—including genomics, transcriptomics, proteomics, metabolomics, and interact omics—to identify promising therapeutic targets with higher accuracy and speed.

Table 3: Comprehensive Overview of AI in Target Identification & Validation

Category	Detailed Description
Objective	The goal of validation and target identification is to pinpoint biologically relevant molecules—typically proteins, genes, or non-coding RNAs—that are involved in disease pathways and are amenable to therapeutic modulation. AI enhances this process by mining vast biological datasets, integrating omics layers (genomics, transcriptomics, proteomics), and modeling biological networks to identify novel, high-confidence drug targets.
Key AI Techniques	- *Machine Learning (ML):* Supervised and unsupervised models analyze multi-dimensional datasets (e.g., gene expression, mutation frequencies) to classify potential targets based on disease relevance. - *Natural Language Processing (NLP):* Extracts knowledge from millions of biomedical documents (PubMed, patents, clinical trials) to detect implicit gene-disease relationships, using named entity recognition (NER) and relation extraction algorithms. - *Graph Neural Networks (GNNs):* Capture intricate relationships in protein–protein interaction (PPI) and gene regulatory networks. These models learn embeddings of network components to predict unknown interactions. - *Deep Learning (DL):* Autoencoders and convolutional networks integrate and reduce dimensionality of omics datasets (RNA-seq, ChIP-seq) to uncover latent patterns relevant to target biology. - *Knowledge Graphs (KGs):* Structure biological knowledge (e.g., disease-gene-pathway-drug linkages) into graph-based systems where AI infers novel associations using link prediction.
Emerging Pharmacy Software/Platforms	- *IBM Watson Discovery:* Uses NLP and cognitive computing to extract target-related evidence from biomedical text. - *Open Targets Platform:* A collaborative public–private partnership between EMBL-EBI, GSK, Biogen, and others; integrates genetics, expression, and literature data to rank potential drug targets by AI-scored evidence. - *Pharos (NIH/IDG):* Provides AI-curated target development levels (TDLs), integrating expression, structure, and binding data for over 20,000 human proteins. - *BioXpress:* A cancer-specific gene expression database powered by AI algorithms to identify over- or under-expressed genes from RNA-seq datasets. - *DeepTarget:* Utilizes transcriptomic and epigenetic features with deep learning to classify gene targets based on tissue-specific and cancer-relevant biomarkers. - *TargetMine:* Integrates functional genomics and protein information; uses AI for scoring and prioritization.
Recent Technological Advancements	- *Multi-Omics Integration with AI:* Tools like DeepOmix use variationalautoencoders to combine DNA methylation, proteomics, and transcriptomics data to discover central regulatory targets in diseases like cancer and neurodegeneration ([Vamathevan et al., 2019]). - *CRISPR Screening Enhanced by AI:* Deep learning interprets CRISPR-Cas9 knockout data to reveal genes essential for cell survival in disease contexts ([Carvalho et al., 2021]). - *Target Druggability Prediction:* Tools like TargetDB use ML models trained on known drug-target pairs to estimate the druggability of novel genes using sequence and structure features. - *Literature-Based Target Extraction:* NLP models like BERT and BioBERT extract disease-gene-drug triads from full-text articles, speeding up evidence-based hypothesis generation ([Lee et al., 2020]). - *AI for Undruggable Proteins:* Tools such as AlphaFold2 + DeepSite identify cryptic or allosteric binding pockets previously deemed inaccessible, opening new possibilities for target validation.
Advantages of AI Integration	- *Data-Driven Discovery:* AI leverages vast heterogeneous data (clinical, experimental, literature) to uncover hidden insights. - *High Throughput and Scale:* Enables simultaneous evaluation of thousands of potential targets across hundreds of disease states. - *Precision in Target Selection:* Prioritizes high-efficacy, low-toxicity targets using AI-based filtering. - *Personalized Targeting:* AI can stratify targets based on patient subgroups, making way for personalized therapies.
Limitations and Current Challenges	- *Data Quality and Integration Issues:* Inconsistent annotations, missing data, and batch effects hamper model training. - *Black-Box Models:* Deep learning approaches, while powerful, lack interpretability, which hinders regulatory acceptance. - *Validation Bottleneck:* Predicted targets require expensive and time-consuming wet-lab validation, slowing translation. - *Biological Complexity:* AI may struggle to fully capture nonlinear, context-specific biological interactions like feedback loops and epigenetic regulation. - *Regulatory & Ethical Concerns:* Use of AI in high-stakes decisions (e.g., oncology targets) needs ethical oversight and explainability frameworks.

a) Machine Learning for Multi-Omics Integration

Machine learning approaches like random forests (RF), ensemble learning models, and support vector machines (SVM) are frequently employed to classify disease-relevant genes or proteins from massive multi-omics datasets. These methods can extract features, cluster disease subtypes, and prioritize potential drug targets based on their functional roles and network topologies [1,2].

For example, tools like *DeepTarget* leverage neural networks to combine gene expression, mutation frequency, and pathway involvement to predict viable targets in cancer therapy (Zeng et al., 2020) [3]. Similarly, *PANDAomics* by Insilico Medicine utilizes AI to rank targets based on disease association, biological relevance, druggability, and novelty [4].

b) Protein-Protein Interaction (PPI) Networks and AI

Deep learning has shown promise in modeling protein-protein interactions, a crucial aspect of identifying nodes central to disease progression. Platforms such as *STRING, BioGRID*, and *HINT* offer curated PPI databases that, when coupled with graph neural networks (GNNs), can reveal hidden relationships within the proteome. These insights enable the identification of key regulatory proteins and interaction hubs [5,6].

c) Predictive Modeling and Literature Mining

Models for natural language processing (NLP), among them Bidirectional Encoder Representations from Transformers (BERT)and SciBERT, are applied to mine biomedical literature and databases like PubMed, identifying emerging targets, associated pathways, and biomarkers [7]. AI-driven tools like IBM Watson Discovery can analyze thousands of scientific papers to extract meaningful patterns and hypotheses [8].

d) Structural and Functional Annotation with AI

AI also contributes to functional annotation and structural prediction of potential targets. AlphaFold, developed by DeepMind, revolutionized protein structure prediction with over 90% accuracy, enabling the visualization of binding sites and aiding in structure-based drug design. [9]

Case Example: A study by Aliper et al. (2016) demonstrated the use of DL models trained on transcriptomic profiles to distinguish between cancerous and non-cancerous cells and identify differential gene expression patterns that could serve as target leads for specific cancer types. These models significantly outperformed conventional clustering techniques [10].

Overall, AI empowers researchers to overcome data complexity and variability in biological systems, enhancing the efficiency and accuracy of target identification and validation. It paves the way for personalized therapeutic strategies by identifying targets that are context-specific and more likely to succeed in downstream development stages.

VI. AI in Virtual Screening and Drug Design

Virtual screening (VS) and drug design are critical components of the early drug discovery process. These techniques aim to identify potential lead compounds with high affinity for a target protein by screening large chemical libraries. Traditional approaches include ligand-based and structure-based virtual screening using molecular docking and pharmacophore modeling. However, these methods often suffer from limited accuracy, high false-positive rates, and significant computational burden.

Artificial Intelligence (AI) addresses these challenges by enabling more precise predictions of molecular interactions, binding affinities, and drug-likeness properties. With advancements in deep learning, generative modeling, and reinforcement learning, AI has transformed both virtual screening and de novo drug design.

a) AI-Driven Virtual Screening

Deep learning models such as convolutional neural networks (CNNs), graph neural networks (GNNs), and recurrent neural networks (RNNs) are employed to predict the bioactivity of compounds against a given target. Tools like *AtomNet* utilize 3D CNNs for structure-based virtual screening, enabling accurate identification of active compounds [11]. *DeepChem* and *Chemprop* are other open-source platforms that provide ML-based frameworks for property prediction, binding affinity estimation, and molecular classification [12,13].

Graph-based deep learning models excel at representing molecular structures and interactions. GNNs consider atom and bond features as nodes and edges, respectively, and can predict activity, toxicity, and solubility with high accuracy. These models significantly outperform traditional quantitative structure-activity relationship (QSAR) methods.

b) AI in De Novo Drug Design

AI models like Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Reinforcement Learning (RL) are increasingly used to generate novel chemical structures with desired properties. The generative model learns the chemical space and generates synthetically feasible molecules optimized for drug-likeness, ADMET properties, and target binding.

One prominent example is *Insilico Medicine*, which designed and synthesized potent DDR1 kinase inhibitors using a generative pipeline combining RL and GANs in less than 18 months—a process that traditionally takes 4–6 years [14]. Similarly, BenevolentAI and Exscientia use AI for automated compound generation and optimization, achieving high hit-to-lead ratios.

c) AI for Docking and Binding Affinity Prediction

AI-powered docking algorithms use DL models to predict ligand-protein binding poses and scoring functions. Tools like *DeepDock, OnionNet,* and *KDEEP* have been shown to outperform classical scoring functions in blind docking challenges [15,16]. These models learn spatial features from protein-ligand complexes and generalize across diverse targets.

d) Case Studies and Applications

- *COVID-19 Drug Discovery:* AI-based drug screening was rapidly employed to identify potential inhibitors of SARS-CoV-2 proteins. Benevolent AI identified baricitinib as a repurposing candidate for COVID-19 treatment, later validated in clinical settings [17].

- *AI-Based Fragment Screening:* Schrödinger's Glide and DeepDock tools were used in campaigns for oncology and CNS diseases, integrating fragment-based drug discovery (FBDD) with DL models to improve early-stage hit generation.

AI's role in virtual screening and molecular design is revolutionizing lead identification and optimization. It offers unprecedented scalability, adaptability, and predictive accuracy, reducing time-to-hit and increasing the probability of downstream success.

VII. AI in Preclinical Testing and Safety Assessment

Preclinical testing is a crucial phase in the drug development process, involving in vitro (cell culture) and in vivo (animal) studies to evaluate the safety, toxicity, pharmacokinetics (PK), and pharmacodynamics (PD) of drug candidates. This phase is vital for understanding how a drug behaves in a biological system and determining whether it is safe enough to progress to human clinical trials. However, traditional preclinical models are costly, time-consuming, ethically controversial, and not always predictive of human outcomes.

Artificial Intelligence (AI) is increasingly being applied to improve the predictive accuracy of preclinical assessments and reduce dependence on animal testing. AI enables the analysis of large-scale toxicogenomic, pharmacogenomic, and bioassay data to forecast potential adverse events, optimize compound dosing, and model drug metabolism.

a) In Silico Toxicity Prediction

Machine learning models such as random forests, support vector machines (SVM), and deep neural networks (DNN) are widely used to predict toxicity endpoints including hepatotoxicity, cardiotoxicity, genotoxicity, and nephrotoxicity. Tools such as *DeepTox, ProTox-II*, and *ADMETlab* apply AI to analyze chemical structures and predict toxicological outcomes before any laboratory testing [18,19].

b) Predictive Pharmacokinetics and Metabolism

AI models can simulate Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) profiles of drug candidates. For instance, the *pkCSM* tool uses graph-based signatures to forecast oral bioavailability, blood-brain barrier penetration, and cytochrome P450 interactions [20]. Deep learning approaches like those used by *ADMET Predictor* (Simulations Plus) further enhance prediction accuracy and guide compound optimization.

c) Systems Biology and Organs-on-Chips

Systems biology modeling, powered by AI, integrates omics and physiological data to simulate tissue-level responses to drugs. When combined with microfluidic technologies (organ-on-chip), AI can model complex biological interactions to predict human-relevant outcomes with greater reliability. For example, AI-enhanced liver-on-chip models can detect hepatotoxicity more accurately than standard animal models [21].

d) Drug-Drug Interaction (DDI) Prediction

AI also plays a vital role in predicting potential drug-drug interactions, a common reason for post-marketing drug withdrawals. Deep learning models trained on electronic health records (EHRs), pharmacovigilance databases, and molecular data can anticipate adverse interactions early in the pipeline. *DeepDDI* is one such model using deep neural networks to identify clinically relevant DDIs [22].

e) Case Studies

- *Merck* has integrated AI models with high-content screening data to predict neurotoxicity and prioritize safe leads earlier in the process.
- *Novartis* and *Atomwise* collaborate to use AI in predicting mitochondrial toxicity, one of the leading causes of late-stage failure.

Figure 3: AI in Preclinical Safety Workflow (A flowchart depicting data ingestion → feature extraction → toxicity/PK/PD modeling → output visualization and risk scoring.)

By leveraging AI in preclinical testing, pharmaceutical companies can identify safety liabilities earlier, minimize reliance on animal models, and accelerate regulatory submissions with higher confidence.

VIII. AI in Clinical Trials and Patient Stratification

Clinical trials are essential for evaluating the safety and efficacy of new drugs in humans, but they are also among the most expensive and time-consuming phases of drug development. Traditional trial designs often suffer from high failure rates, low recruitment efficiency, and lack of personalized treatment strategies. Artificial Intelligence (AI) offers novel solutions to these challenges by enhancing patient recruitment, optimizing trial design, stratifying patients, and enabling real-time monitoring.

a) Patient Recruitment and Matching

Recruiting eligible participants is a major bottleneck in clinical trials. AI-driven natural language processing (NLP) systems can analyze electronic health records (EHRs), clinical notes, and diagnostic reports to match patients with trial inclusion and exclusion criteria. For instance, IBM Watson for Clinical Trial Matching has shown a 70% reduction in trial screening time by

automatically identifying qualified participants from clinical databases [23].

b) Trial Design Optimization

Machine learning models are used to simulate various trial scenarios, allowing researchers to predict potential outcomes and adapt protocols accordingly. Bayesian adaptive trial designs, supported by AI, can adjust randomization probabilities based on interim results, improving trial efficiency and ethical considerations. AI also helps in selecting endpoints, optimizing dosing regimens, and predicting dropout rates [24].

c) Patient Stratification and Precision Medicine

AI can segment patients into subgroups based on biomarkers, genetic profiles, lifestyle data, and disease phenotypes. Unsupervised learning algorithms like clustering and t-SNE are used to identify hidden patient subpopulations that may respond differently to treatments. This stratification supports precision medicine, ensuring the right drug is given to the right patient at the right time [25].

d) Real-Time Monitoring and Remote Data Collection

With the rise of wearable devices and digital health platforms, AI facilitates continuous remote monitoring of trial participants. Deep learning algorithms can analyze data from heart rate monitors, glucose sensors, and sleep trackers to detect early signs of adverse events or non-compliance. These tools enhance patient safety and reduce the need for in-person visits [26].

e) Case Study: AI in Oncology Trials

Pfizer partnered with Concerto HealthAI to use real-world evidence and AI to optimize oncology trials. The collaboration helped improve trial feasibility, identify responsive patient groups, and reduce protocol amendments, which are typically time-consuming and costly.

By integrating AI into clinical trials, the pharmaceutical industry is moving toward more dynamic, data-driven, and patient-centric research models. These innovations can shorten trial durations, improve outcome predictability, and increase regulatory acceptance of trial results.

IX. Challenges, Limitations, and Ethical Considerations

While the integration of Artificial Intelligence (AI) into drug discovery and development offers remarkable potential, several challenges and ethical concerns must be addressed to ensure responsible and effective implementation.

a) Data Quality and Standardization

AI models rely heavily on high-quality, well-annotated datasets. Inconsistent data formatting, missing values, and heterogeneous sources (e.g., clinical, genomic, imaging) can reduce model accuracy and reproducibility. Moreover, lack of standardization in data collection protocols across institutions hinders the integration and generalization of AI models [27].

b) Interpretability and Transparency

Most deep learning models function as "black boxes," making it difficult to interpret how specific decisions are made. This lack of transparency can hinder the trust of regulatory bodies, clinicians, and patients. Developing explainable AI (XAI) models is crucial to ensure traceability, accountability, and informed decision-making in healthcare applications [28].

c) Regulatory and Validation Frameworks

AI-based tools used in drug discovery must adhere to stringent validation standards before regulatory acceptance. However, there is a lack of clear regulatory guidelines tailored specifically for AI systems. The U.S. FDA and European Medicines Agency (EMA) are actively developing frameworks, but harmonization across global agencies remains a challenge [29].

d) Bias and Generalizability

AI models can inadvertently reflect biases present in training datasets, leading to unequal performance across different patient populations. This is especially concerning in precision medicine, where biased models can result in suboptimal or unsafe treatment recommendations for underrepresented groups [30].

e) Ethical and Privacy Concerns

AI applications in healthcare often require access to sensitive patient data. Ensuring patient privacy, data ownership, and compliance with regulations such as the General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA) is essential. Additionally, ethical dilemmas arise when AI systems are involved in life-altering decisions without sufficient human oversight [31].

f) Talent and Infrastructure Gaps

The implementation of AI technologies in drug development requires a skilled workforce proficient in data science, biology, and regulatory science. Many pharmaceutical companies face challenges in building interdisciplinary teams and developing the necessary computational infrastructure to support AI workflows [32].

g) Cost and Resource Allocation

Although AI promises long-term savings, its initial implementation can be costly. Investments are needed in data infrastructure, high-performance computing, model training, and integration into existing R&D pipelines. For smaller firms and academic

institutions, such costs may be prohibitive without collaborative partnerships.

Despite these challenges, continued innovation, policy development, and interdisciplinary collaboration can help address limitations and foster responsible AI use in drug discovery.

X. Future Directions and Conclusion

The application of Artificial Intelligence (AI) in drug discovery and development is rapidly evolving, offering transformative potential across the entire pharmaceutical value chain. Looking ahead, several promising directions are expected to redefine the landscape of biomedical research and therapeutic innovation.

The future of AI in drug discovery and development holds tremendous promise. Integration of multi-omics data-combining genomic, proteomic, metabolomic, and other biological information-will enable more comprehensive modeling of disease mechanisms and drug effects. This holistic approach will allow for more precise target identification and personalized treatment strategies.

Advances in quantum computing may further accelerate AI applications in drug discovery by enabling more complex molecular simulations and property predictions. Federated learning approaches could facilitate collaborative research while preserving data privacy, allowing organizations to collectively train AI models without sharing sensitive information.

The combination of AI with laboratory automation represents another promising direction. AI-guided robotic systems can design, execute, and analyze experiments with minimal human intervention, creating a closed-loop discovery process that iteratively improves based on experimental results.

As AI technologies continue to evolve, their integration into the drug development pipeline will likely become more seamless and comprehensive. This evolution will require ongoing collaboration between AI researchers, drug developers, clinicians, and regulatory authorities to ensure that AI-driven approaches deliver safe, effective, and accessible therapeutic innovations.

a) Integration with Multi-Modal Data

The future of AI in drug discovery lies in the integration of diverse data modalities, including genomics, proteomics, transcriptomics, metabolomics, imaging, and electronic health records (EHRs). Multi-modal AI models will enable a more holistic understanding of disease mechanisms and drug responses, allowing for precise, patient-specific therapeutic interventions. Tools like Google's DeepMind and Meta AI are advancing this capability with deep learning architectures capable of processing heterogeneous biomedical data at scale.

b) Federated and Privacy-Preserving Learning

Federated learning models are gaining attention for their ability to train AI algorithms across decentralized datasets without sharing sensitive patient data. This approach helps overcome data privacy concerns and facilitates cross-institutional collaborations. As privacy regulations become more stringent, such models will play a critical role in clinical AI applications.

c) AI-Augmented Drug Repurposing

AI is expected to significantly enhance drug repurposing strategies by identifying new therapeutic uses for existing drugs. This is particularly valuable in pandemic scenarios and rare diseases, where time and resources are limited. Platforms like BenevolentAI and Healx are already pioneering this field with AI models trained on biomedical knowledge graphs and real-world evidence.

d) Digital Twins and Personalized Drug Testing

The concept of digital twins—virtual replicas of individual patients—combined with AI could revolutionize personalized medicine. These models simulate disease progression and drug response in silico, enabling personalized treatment regimens, dosing strategies, and adverse event prediction before administration.

e) Regulatory Harmonization and Ethical AI

Future success of AI in drug development will depend on the establishment of globally harmonized regulatory frameworks that address data integrity, model validation, bias mitigation, and ethical considerations. Collaboration between regulators, industry stakeholders, and academic institutions will be essential to achieve trust and widespread adoption.

f) Enhanced Human-AI Collaboration

Rather than replacing scientists, AI will increasingly serve as a decision-support partner, augmenting human capabilities. The synergy between domain experts and AI tools will accelerate hypothesis generation, streamline experimentation, and improve R&D productivity.

XI. Conclusion

AI has emerged as a powerful force in transforming drug discovery and development, from target identification to clinical trials. By enabling faster, cheaper, and more accurate research processes, AI holds the promise of accelerating the delivery of safer and more effective therapies to patients. Despite existing challenges related to data quality, model transparency, and regulatory readiness, the future is optimistic. As technologies mature and collaborative ecosystems evolve, AI will become an indispensable pillar of precision medicine and pharmaceutical innovation.

Artificial intelligence has emerged as a transformative force in drug discovery and development, offering innovative solutions to longstanding challenges in pharmaceutical research. By accelerating target identification, enabling virtual screening of vast compound libraries, predicting molecular properties, optimizing clinical trials, and advancing personalized medicine, AI technologies are reshaping the entire drug development landscape.

While challenges related to data quality, regulatory considerations, and ethical concerns persist, the rapid pace of innovation in AI methodologies suggests that many of these limitations will be addressed in the coming years. The integration of diverse AI techniques-from machine learning and deep learning to generative AI and large language models-provides a rich toolkit for addressing complex problems across the drug development pipeline.

The successful implementation of AI in pharmaceutical research requires interdisciplinary collaboration, combining expertise in computer science, biology, chemistry, medicine, and regulatory affairs[1]. By fostering such collaborative approaches and continuing to advance AI technologies, the field stands poised to dramatically accelerate the discovery and development of novel therapies, potentially transforming patient care and reducing the global burden of disease.

As we look to the future, the synergy between human expertise and artificial intelligence will likely define a new paradigm for drug discovery and development-one characterized by greater efficiency, reduced costs, higher success rates, and more personalized therapeutic approaches. This evolution promises to benefit not only the pharmaceutical industry but also healthcare systems and, most importantly, patients awaiting new treatments for challenging medical conditions.

References Références Referencias

1. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics Data Integration, Interpretation, and Its Application. BioinformBiol Insights. 2020;14: 1177932219899051.
2. Angermueller C, Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. MolSyst Biol. 2016;12(7):878.
3. Zeng X, Song X, Ma T, Pan X, Zhou Y, Hou Y, Zhang Z. DeepTarget: End-to-end learning framework for predicting protein targets of small molecules. J Cheminform. 2020;12(1):49.
4. Zhavoronkov A, Vanhaelen Q, Oprea TI. Will Artificial Intelligence for Drug Discovery Impact Clinical Pharmacology? Clin Pharmacol Ther. 2020;107(4): 780–785.
5. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–D613.
6. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. Nucleic Acids Res. 2006;34 (Database issue):D535–D539.
7. Beltagy I, Lo K, Cohan A. SciBERT: A Pretrained Language Model for Scientific Text. arXiv preprint arXiv:1903.10676. 2019.
8. Ferrucci D. Introduction to "This is Watson". IBM J Res Dev. 2012;56(3.4):1:1–1:15.
9. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with Alpha Fold. Nature. 2021;596(7873):583–589.
10. Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. Mol Pharm. 2016;13(7):2524–2530.
11. Wallach I, Dzamba M, Heifets A. Atom Net: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. arXiv preprint arXiv:1510.02855. 2015.
12. Ramsundar B, Eastman P, Walters P, Pande V. Deep Learning for the Life Sciences. O'Reilly Media; 2019.
13. Yang K, Swanson K, Jin W, Coley CW, Eiden P, Gao H, et al. Analyzing Learned Molecular Representations for Property Prediction. J Chem Inf Model. 2019;59(8):3370–3388.
14. Zhavoronkov A, Ivanenkov YA, Aliper A, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol. 2019;37(9):1038–1040.
15. Jiménez J, Skalic M, Martinez-Rosell G, De Fabritiis G. KDEEP: Protein–ligand absolute binding affinity prediction via 3D-convolutional neural networks. J Chem Inf Model. 2018;58(2):287–296.
16. Liu Z, Li Y, Han L, Li J, Liu J, Zhao Z, et al. PDB-wide collection of binding data: current status of the PDBbind database. Bioinformatics. 2015;31(3):405–412.
17. Richardson P, Griffin I, Tucker C, Smith D, Oechsle O, Phelan A, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. Lancet. 2020;395(10223):e30–e31.
18. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: Toxicity prediction using deep learning. Front Environ Sci. 2016;3:80.
19. Banerjee P, Eckert AO, Schrey AK, Preissner R. ProTox-II: A webserver for the prediction of toxicity of chemicals. Nucleic Acids Res. 2018;46 (W1):W257–W263.
20. Pires DE, Blundell TL, Ascher DB. pkCSM: Predicting small-molecule pharmacokinetic and

toxicity properties using graph-based signatures. J Med Chem. 2015;58(9):4066–4072.
21. Zhang B, Korolj A, Lai BFL, Radisic M. Advances in organ-on-a-chip engineering. Nat Rev Mater. 2018;3(8):257–278.
22. Ryu JY, Kim HU, Lee SY. Deep learning improves prediction of drug–drug and drug–food interactions. ProcNatlAcadSci USA. 2018;115(18):E4304–E4311.
23. Ross C, Swetlitz I. IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show. STAT News. 2018.
24. Bhatt DL, Mehta C. Adaptive Designs for Clinical Trials. N Engl J Med. 2016;375(1):65–74.
25. Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J, et al. Artificial intelligence and machine learning in clinical development: a translational perspective. NPJ Digit Med. 2019;2:69.
26. Montalbano A, Toumazou C. Wearable sensors for real-time clinical trial monitoring: current trends and future opportunities. ACS Sens. 2020;5(7):2244–2260.
27. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA. 2018;319(13):1317–1318.
28. Holzinger A, Biemann C, Pattichis CS, Kell DB. What do we need to build explainable AI systems for the medical domain? Rev Life Sci. 2017;36:6–10.
29. U.S. Food & Drug Administration. Artificial Intelligence and Machine Learning in Software as a Medical Device. 2021.
30. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. JAMA Intern Med. 2018;178(11):1544–1547.
31. Price WN, Cohen IG. Privacy in the age of medical big data. Nat Med. 2019;25(1):37–43.
32. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56.
33. IBM Watson Health, Mayo Clinic. AI to accelerate patient matching. [Online]. Available: https://www.ibm.com/watson-health/
34. FDA. Artificial Intelligence/Machine Learning-Based Software as a Medical Device (SaMD) Action Plan. 2021.
35. EMA. Big Data Task Force. European Medicines Agency, 2020.
36. Insilico Medicine. AI-designed drug enters clinical trials. [Press Release]. 2021.
37. Artificial intelligence in drug discovery and development. PubMed Central (PMC). Published October 21, 2020. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC7577280/
38. Artificial Intelligence in Pharmaceutical Technology and Drug Development. PubMed Central (PMC). Published July 10, 2023. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10385763/
39. Early evidence and emerging trends: how AI is shaping drug discovery and clinical development. Drug Target Review. Published March 26, 2025. Available from: https://www.drugtargetreview.com/article/158593/early-evidence-and-emerging-trends-how-ai-is-shaping-drug-discovery-and-clinical-development/
40. Artificial intelligence in drug discovery and development. In: Artificial Intelligence in Drug Discovery and Development. ScienceDirect. Available from: https://www.sciencedirect.com/science/article/abs/pii/B9780443223082000184
41. Artificial intelligence for drug discovery and development in neurodegenerative diseases. Science Direct. Available from: https://www.sciencedirect.com/science/article/pii/S0959440X24000034
42. Artificial intelligence in drug discovery and development. PubMed. Available from: https://pubmed.ncbi.nlm.nih.gov/33099022/
43. Artificial intelligence in drug development: reshaping the therapeutic landscape. Therapeutic Advances in Drug Safety. Published February 13, 2025. Available from: https://journals.sagepub.com/doi/full/10.1177/20420986251321704
44. Nelson JL, Smith BL, Jared JD, Younger JG. Prospective trial of real-time electronic surveillance to expedite early care of severe sepsis. Ann Emerg Med. 2011;57:500–504.
45. Tan C, Chen H, Xia C. The prediction of cardiovascular disease based on trace element contents in hair and a classifier of boosting decision stumps. Biol Trace Elem Res. 2008;129:9–19.
46. Chen H, Tan C. Prediction of Type-2 diabetes based on several element levels in blood and chemometrics. Biol Trace Elem Res. 2011;147:67–74.
47. Hooper MH, Weavind L, Wheeler AP, et al. Randomized trial of automated, electronic monitoring to facilitate early detection of sepsis in the intensive care unit. Crit Care Med. 2012;40:2096–2101.
48. Giannini HM, Ginestra JC, Chivers C, et al. A machine learning algorithm to predict severe sepsis and septic shock: development, implementation, and impact on clinical practice. Crit Care Med. 2019.
49. Naushad SM, Janaki Ramaiah M, Pavithrakumari M, et al. Artificial neural network-based exploration of gene-nutrient interactions in folate and xeno-biotic metabolic pathways that modulate susceptibility to breast cancer. Gene. 2016;580:159–168.
50. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463–477.

51. Gawehn E, Hiss JA, Schneider G. Deep learning in drug discovery. Mol Inform. 2016;35(1):3–14.
52. Gómez-Bombarelli R, Wei JN, Duvenaud D, et al. Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci. 2018;4(2):268–276.
53. Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018;17(2):97–113.
54. Miotto R, Li L, Kidd BA, Dudley JT. Deep Patient: an unsupervised representation to predict the future of patients from the electronic health records. Sci Rep. 2016;6:26094.
55. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J ClinPharmacol. 1998;54 (4):315–321.
56. Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. Clin Pharmacol Ther. 2018;104(1):42–52.
57. Wang Y, Zeng D, Wang L. Social media analytics for drug safety signal detection. ACM Trans Manag Inf Syst. 2018;9(4):1–19.
58. Harpaz R, Callahan A, Tamang S, et al. Text mining for adverse drug events: the promise, challenges, and state of the art. Drug Saf. 2014;37(10):777–790.
59. Botsis T, Nguyen MD, Woo EJ, et al. Text mining for the biopharmaceutical domain: a systematic review of recent advances. Drug Saf. 2011;34(10):809–823.
60. Luo H, Wang J, Li M, et al. Drug repositioning based on comprehensive similarity measures and bi-random walk algorithm. Bioinformatics. 2016; 32(17):2664–2671.
61. Carvalho C, et al. Functional genomics and CRISPR screens in drug discovery: Synergy with AI. Trends Pharmacol Sci. 2021;42(10):765–781.
62. Lee J, et al. BioBERT: A pre-trained biomedical language representation model for biomedical text mining. Bioinformatics. 2020;36(4):1234–1240.
63. Zhou Y, et al. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. Cell Discov. 2020;6:14.
64. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018;23(6):1241–1250.
65. Ekins S, Puhl AC, Zorn KM, et al. Exploiting machine learning for end-to-end drug discovery and development. Nat Mater. 2019;18(5):435–441.
66. Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180(4):688–702.e13.
67. Vamathevan J, Clark D, Czodrowski P, et al. Machine learning in drug discovery: progress and prospects. Drug Discov Today. 2019;24(3):589–601.
68. Altae-Tran H, Ramsundar B, Pappu AS, Pande V. Low data drug discovery with one-shot learning. ACS Cent Sci. 2017;3(4):283–293.
69. Walters WP, Barzilay R. Applications of deep learning in molecule generation and molecular property prediction. AccChem Res. 2021;54(2):263–270.
70. Zhavoronkov A. Artificial intelligence for drug discovery, biomarker development, and generation of novel chemistry. Mol Pharm. 2018;15(10):4311–4313.
71. Ragoza M, Hochuli J, Idrobo E, Sunseri J, Koes DR. Protein–ligand scoring with convolutional neural networks. J Chem Inf Model. 2017;57(4):942–957.
72. Kuenemann M, Schmidl S, Korotcov A, et al. Evaluation of graph neural networks for predicting drug–target interactions. J Chem Inf Model. 2021;61(12):5858–5868.
73. Chen L, Cruz A, Ramsey S, et al. The rise of generative models in drug design. Drug Discov Today. 2020;25(6):1034–1041.
74. Liu Q, Wu H, Wang X, et al. Application of artificial intelligence in the diagnosis and treatment of COVID-19: A systematic review. Int J Med Sci. 2021;18(3):678–689.
75. Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V. Deep neural nets as a method for quantitative structure–activity relationships. J Chem Inf Model. 2015;55(2):263–274.
76. Elton DC, Boukouvalas Z, Fuge MD, Chung PW. Deep learning for molecular design—a review of the state of the art. MolSyst Des Eng. 2019;4(4):828–849.
77. Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. Proceedings of the 34th International Conference on Machine Learning. 2017:1263–1272.
78. Schwaller P, Gaudin T, Lányi D, Bekas C, Laino T. "Found in translation": predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models. Chem Sci. 2018;9(28):6091–6098.