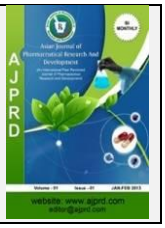
Available online on 15.04.2026 at http://ajprd.com

Asian Journal of Pharmaceutical Research and Development

Open Access to Pharmaceutical and Medical Research



Open Access

Review Article

Revolutionary Aspects of Artificial Intelligence in Drug Discovery and Development

Neetu Dharu N, Iram Fatema Siddique, Dr. Gautam B. Sonara, Dr. Manish B. Goyani

Vidhyadeep College of pharmacy, Surat,Vidhyadeep University, Surat

A B S T R A C T

The pharmaceutical sector is shifting with the help of artificial intelligence (AI), which has enhanced the accuracy of drug development and discovery. AI enables predictive modeling of bioactivity and toxicity, accelerates virtual screening and lead optimization, enhances the processing of complex biological data sets, and streamlines clinical trial design and patient recruitment. To analyze how AI has transformed the pharmaceutical sector, enhanced product life cycle management, and facilitated data-informed decision-making from target candidate selection to market, this research includes high-impact studies. We also address operational, ethical, and regulatory challenges as well as what needs to be done to ensure that AI delivers its promise of creating better, faster, and cheaper cures.

Keywords: Artificial intelligence, Drug discovery, Deep learning, Bioactivity prediction, Toxicity prediction, pharmaceutical industry transformation

A R T I C L E I N F O: Received 10 Jan 2026; Review Complete 28 Feb. 2026; Accepted 20 March 2026; Available online 15 April. 2026

Cite this article as:

Dharu N, Iram Fatema Siddique IF, Dr. Gautam B. Sonara, Dr. Manish B. Goyani, Revolutionary Aspects of Artificial Intelligence in Drug Discovery and Development, Asian Journal of Pharmaceutical Research and Development. 2026; 14(2):51-61

DOI: http://dx.doi.org/10.22270/ajprd.v14i2.1720

*Address for Correspondence:

Neetu Dharu, Vidhyadeep College of pharmacy, Surat, Vidhyadeep University, Surat

INTRODUCTION

Drug development and discovery are among the most time-consuming, complicated, and expensive biomedical research processes. It took historically 10 to 15 years and more than USD 2 to 3 billion to develop a new medicinal compound. Moreover, the rate of attrition is so high that fewer than 10% of the ones that reach the clinic actually receive regulatory approval. The time-consuming process is often responsible for the delayed dispensing of life-saving drugs and is a key factor in the outrageously high cost of healthcare globally. The pharmaceutical sector has been seeking new ways to make drug discovery better, more precise, and cost-effective for decades; over the past decade or so, artificial intelligence (AI) has been a game-changer [1].

Due to rapid developments in machine learning (ML), deep learning (DL), and big data analytics, artificial intelligence (AI) or the imitation of human intelligence by computational models that learn, reason, and make decisions has also evolved rapidly. AI facilitates the handling of large and heterogeneous datasets in pharmaceutical research and development, such as genomics, proteomics, cheminformatics, and clinical data, to build predictive models capable of discovering new therapeutic targets, designing promising drug molecules, and optimizing their properties with unprecedented efficiency and precision. AI-driven methods allow for in silico prediction and hypothesis generation, which is less expensive and time-consuming compared to traditional methods based primarily on trial-and-error and extensive laboratory testing[2].

The ability of AI to enhance acknowledgement hit recognition and lead promotion is among its most ground-breaking accomplishment. Researchers can choose only the denotes favourable candidates for 14 experimental and verifiying by using deep learning algorithms to analyze millions of chemical structures and for determining bioactivity, pharmacokinetics, and toxicity profiles(3,4,8). AI can helps in renew of drugs, which include methodically screening existing

components for new therapeutic properties. This procedure has also become more common amid worldwide health crises such as the COVID-19 pandemic[3].

AI plays important role in life cycle of drug management, from preclinical studies to clinical trial design, patient stratification, and post-market surveillance, in addition to drug development.AI-driven foretell analytics can identify high-risk adverse drug reactions, maximize dosing regimens and monitor real-world drug performance, thereby supporting accuracy medicine and improvement patient safety.

Consequently, contract research organizations (CROs), pharmaceutical corporations, and regulatory bodies are rapid investing at AI-powered platforms to gain competitively benificial, reduce risk, and accelerate inventation[4].

This review explores the impact of AI in drug discovery and development, discussing how it has changed the pharmaceutical industry, improved the product lifecycle, supported drug management, and revised medication development. AI applications in drug screening, bioactivity prediction, and toxicity assessment, highlighting their role in improving efficiency, reducing failure rates, and enabling the development of safer and more effective therapeutics[5].

RECENT STATISTICS & MARKET TRENDS [6]

1. Market Size & Growth

- It is forecast to reach USD 7.1 billion by 2030, growing at a compound annual growth rate (CAGR) of ~23.7%.
- Another study estimates that in 2023 the market was around USD 1.48billion, with projected growth to about USD 15.5 billion by 2032, at a CAGR of ~29.9%.
- The AI in pharmaceutical market more broadly (beyond just discovery) is also growing fast: for example, it was USD 1.53 billion in 2024, expected to rise to USD 22.90 billion by 2034, with a CAGR of ~29.8%.

2. Generative AI Sub-Segment

The generative AI (8,14,15) in drug discovery segment was valued at about USD 204.56 million in 2024, expected to reach roughly USD 2,300.52 million (~USD 2.3 billion) by 2034 (CAGR ~27.4%).

3. Applications & Adoption Trends

- Drug optimization & repurposing accounted for a large share (~40%) of AI applications in drug discovery in 2023.
- Machine learning still leads in technology share (~38.8% in 2024) among AI technologies in pharmaceuticals; generative AI (8,14,15) is growing more rapidly though, projected to expand at a higher CAGR.
- Regionally, **North America** holds the largest share currently; however**Asia-Pacific** is among the fastest growing.

4. Impact Metrics

- AI is seen to reduce drug discovery & development timelines significantly, and to lower costs. Some reports estimate up to halving R&D timelines or delivering candidates faster. (Exact numbers vary depending on the phase and application)
- The success rate for drug candidate generation and early preclinical-phase selection is improving with AI, especially in identifying targets, predicting toxicity (7,9,17), and optimizing pharmacokinetics[7].

Real-World Examples

1. Insilico Medicine: Generative AI Drug Entering Phase II

- A totally generative AI (8,14,15)-designed medication candidate (INS018_055) for idiopathic pulmonary fibrosis (IPF) is available from Insilico Medicine.
- This candidate has moved into Phase II trials in both China and the U.S., following favorable safety, tolerability, and pharmacokinetics results.
- The AI platform “Pharma.AI” from Insilico includes tools such as **PandaOmics** (target discovery), **Chemistry42** (molecule generation), and **InClinico** (clinical trial predictions) [8].

1. Alpha Fold & Structure-Based Discovery

- DeepMind's AlphaFold (1,2) has significantly paved the way for structure-based drug discovery by predicting protein structures, including several for which no experimental structure was known.
- For example, a small molecule inhibitor of Cyclin-dependent Kinase-20 (CDK20), a previously unexperimentally solved structure target, was hit and optimized rapidly (first hit within 30 days, with only ~7 compounds being made; followed by more active hit after second round) with the aid of AlphaFold (1,2) predicted structures and AI tools[9].

2. Corporate & Partnership Moves

- With an investment of more than USD 1 billion, Eli Lilly has unveiled TuneLab, a platform based on machine learning and artificial intelligence (AI/ML) that seeks to provide access to its extensive drug research databases. The goal of the change is to make it possible for smaller biotechnology firms to use advanced computational models for medication development, safety assessment, and candidate identification.
- In order to develop innovative treatments for obesity, type 2 diabetes, and cardiovascular disorders, Novo Nordisk has expanded its

strategic alliance with Valo Health to use artificial intelligence technology in conjunction with human-derived information [10].

How These Strengthen the "Revolutionary" Claim

- Multibillion-dollar markets are expanding at robust double-digit compound annual growth rates (CAGR), demonstrating that artificial intelligence (AI) is not only a far-off promise but is now generating tangible economic value.
- The development of the first drug candidate fully identified by generative AI to Phase II clinical trials, the demonstration of structure prediction tools (e.g., AlphaFold) to predict drug hits without experimental input, and significant investments made by pharmaceutical companies that imply AI is becoming a crucial component of biomedical research are all examples of pivotal milestones[11].
- Notably, compared to traditional discovery techniques, AI-facilitated strategies demonstrate shorter turnaround times and reduced resource consumption, achieved through the synthesis of fewer compounds and streamlined development cycles, signaling a radical shift in pharmaceutical innovation.

1. Transformation of the Pharmaceutical Industry by AI

Throughout history, the pharmaceutical sector has been characterized by long, expensive, and high-risk research and development (R&D) cycles. Developing a single new drug typically requires 10–15 years and USD 2–3 billion, with a clinical success rate of less than 10% for drug candidates entering human trials. This conventional model, largely dependent on empirical experimentation and sequential workflows, has been increasingly challenged by the growing complexity of diseases, rising costs of clinical trials [12], and the demand for personalized medicine. Artificial intelligence (AI) has appeared in this area. as a disruptive force, transforming nearly every facet of the pharmaceutical value chain—from early-stage target identification to late-stage clinical decision-making and post-marketing surveillance[13].

Computer paradigms that replicate human cognition such as reasoning, adaptive learning, and problem-solving are termed as artificial intelligence (AI). Large biomedical data are exploited by major subdomains such as machine learning (ML), deep learning (DL), natural language processing (NLP), and generative AIto detect patterns beyond our understanding, build predictive models, and enhance decision-making processes with high fidelity and efficiency. These methods are revolutionizing data-directed drug design, prediction of pharmacodynamic and pharmacokinetic properties, and chemical space exploration in pharmaceutical sciences[14].

Discovery and design of new drug pipelines is where AI has made the greatest contribution. AI platforms like BenevolentAI, Atomwise, and Insilico Medicine's PandaOmics and Chemistry42 have shown the potential to rank therapeutic targets by integrative analysis of multi-omics data, carry out enormous in silico screening of billions of compounds, and generate de novo molecular structures with optimal pharmacological profiles. In addition, AI shortens research time periods and costs by computationally predicting new therapeutic uses of established drugs, thus facilitating drug repurposing. During the COVID-19 pandemic, AI-based systems swiftly weeded out possible antivirals, demonstrating the potential of such approaches and their capacity to address new global health emergencies [15].

Beyond discovery, AI is revolutionizing clinical development and operations. Predictive analytics enable the selection of optimal patient cohorts, design of adaptive clinical trials, and real-time monitoring of safety signals. These applications reduce trial failures, improve recruitment efficiency, and shorten approval timelines. AI also supports manufacturing optimization, supply chain forecasting, and pharmacovigilance [16], ensuring that once drugs reach the market, they are produced efficiently, distributed effectively, and monitored for safety in diverse populations.

The economic and strategic implications of this transformation are profound. According to market analyses, the AI in drug discovery market, valued at approximately USD 1.3 billion in 2022, is projected to grow at a CAGR exceeding 23–29%, reaching USD 7–15 billion by 2030–2032. This growth reflects both technological maturity and the widespread adoption of AI solutions by prominent pharmaceutical firms, such as Pfizer, Novartis, Roche, and Eli Lilly, who are integrating AI into their R&D pipelines to gain competitive advantage [17].

Crucially, the shift toward AI-driven drug development aligns with the broader movement toward **precision medicine**, where therapies are tailored to an individual's genetic, phenotypic, and environmental profile. By integrating genomics, transcriptomics, proteomics, and real-world evidence, AI enables a more granular understanding of disease mechanisms and treatment response variability, facilitating the evolution developing more secure and efficient treatments [18].

In summary, the pharmaceutical industry is changing its paradigm, driven by AI's ability to transform massive biomedical data into actionable insights, reduce costs, improve productivity, and bring innovative therapies to market faster. The main areas where AI has had the biggest impact will be examined in this section, including **target discovery, lead optimization, drug repurposing, clinical trial design, manufacturing, and post-market monitoring**, highlighting both the opportunities and The difficulties that come with this technological breakthrough[19].

2. New Business Models

The rise of AI has spurred the formation of AI-native biotech companies (e.g., Insilico Medicine (8,14), Exscientia) and partnerships between technology firms and large pharmaceutical companies. These collaborations focus on integrating AI platforms into discovery pipelines, optimizing R&D resources, and shortening development timelines [20].

R & D Workflow Restructuring

AI is now embedded at multiple decision points: target identification, hit discovery, lead optimization, ADMET prediction, and clinical trial planning. This integration enables iterative, data-driven workflows that are more adaptive and efficient. Hybrid human–AI decision-making is increasingly common, where AI-generated hypotheses are validated by experimental scientists[21].

AI-Enhanced Product Lifecycle Management

Artificial Intelligence (AI) is not limited to accelerating drug discovery; its Possibility for transition extends across the entire pharmaceutical product lifecycle, ensuring a seamless transition from concept to commercialization. AI enables a fully integrated, data-driven ecosystem that enhances efficiency, quality, and compliance while reducing overall development timelines and costs. By combining machine learning (ML), deep learning (DL), predictive analytics, and natural language processing (NLP), AI empowers decision-making at every stage of the pharmaceutical value chain[22].

1. Discovery and Early Research

In the earliest stages of drug development, AI addresses one of the most critical bottlenecks: identifying disease-relevant targets and prioritizing compounds with the highest likelihood of success.

- Multi-omics Data Integration: AI platforms such as PandaOmics, DeepMind's AlphaFold, and BenevolentAI integrate genomics, transcriptomics, proteomics, and metabolomics data to map disease pathways, identify novel therapeutic targets, and predict their druggability.
- Virtual Screening & De Novo Design: Deep learning models screen billions of chemical structures *in silico*, performing ligand-based and structure-based virtual screening to rank potential hits. Generative AI tools (e.g., Chemistry42, Molecule.one) design entirely new drug-like molecules that satisfy predefined pharmacological and physicochemical criteria.
- Hit-to-Lead Optimization: AI-driven multi-objective optimization models fine-tune lead compounds by balancing potency, selectivity, solubility, and metabolic stability, minimizing the number of iterative synthesis cycles needed in the lab[23].

2. Preclinical Development

Preclinical studies traditionally consume significant resources and involve extensive in vitro and in vivo testing. AI mitigates these challenges through predictive modeling:

- ADME/Tox Prediction: Predictions of absorption, distribution, metabolism, excretion, and toxicityare made by machine learning algorithms. (ADME/Tox) properties of candidate molecules with high accuracy, reducing reliance on animal testing.
- Mechanistic Toxicology & Off-Target Effects: AI models such as DeepTox or ProTox-II forecast potential organ-specific toxicities, genotoxicity, and off-target interactions early in the pipeline, preventing costly late-stage failures.
- In Silico Pharmacokinetics: Physiologically based pharmacokinetic (PBPK) models simulate drug concentration-time profiles across tissues, guiding dose selection and formulation design before moving to animal studies[24].

3. Clinical Development

AI has a profound impact on clinical trial efficiency, which is one of the most resource-intensive phases of the drug lifecycle.

- Patient Recruitment and Stratification: AI mines electronic health records (EHRs), genomics databases, and social media to identify eligible patients, particularly those with rare diseases, thus reducing recruitment times[25].
- Adaptive Trial Design: Reinforcement learning and Bayesian models support adaptive trial protocols, enabling real-time adjustments in dosing, randomization, and sample size based on interim data, thereby improving trial success rates[26].
- Digital Biomarkers and Remote Monitoring: Wearable sensors and AI-powered digital biomarkers continuously collect patient data, reducing site visits and enhancing compliance monitoring[27],
- Safety Signal Detection: AI-enabled pharmacovigilanceplatforms analyze spontaneous adverse event reports, EHR data, and real-world evidence to identify safety concerns earlier than traditional methods[28].

4. Manufacturing and Quality Control

AI optimizes pharmaceutical manufacturing through smart automation and real-time analytics.

- Process Optimization: AI-based process analytical technology (PAT) monitors critical process parameters (CPPs) and critical quality attributes (CQAs), enabling continuous manufacturing and reducing batch failures.
- Predictive Maintenance: Machine learning predicts equipment failures before they occur, reducing downtime and improving overall equipment effectiveness (OEE).
- Supply Chain and Demand Forecasting: AI models anticipate drug demand, optimize raw material procurement, and manage inventory to avoid shortages or overproduction[29].

5. Commercialization and Post-Market Surveillance

After regulatory approval, AI continues to create value in marketing, pharmacovigilance, and lifecycle management.

- Market Access & Pricing: AI-driven health economic models predict cost-effectiveness, helping manufacturers set competitive yet sustainable prices.
- Personalized Marketing: NLP and predictive analytics analyze prescriber behavior and patient demographics to optimize marketing campaigns and improve therapy adoption.
- Real-World Evidence (RWE) Generation: AI aggregates real-world data from EHRs, claims databases, and patient registries to assess long-term efficacy and safety, supporting label expansion or risk mitigation strategies.
- Adverse Event Prediction: Ongoing surveillance using ML algorithms ensures faster detection of uncommon but serious adverse consequences, enhancing patient safety and regulatory compliance[30].

6. End-to-End Integration and Strategic Benefits

When applied synergistically, these AI applications create a closed-loop product lifecycle management system. Insights generated at later stages (e.g., post-market safety data) are fed back into early discovery pipelines, creating a learning healthcare system that continuously improves future R&D efforts[31].

The benefits include:

- Reduced Overall Cycle Time: Significant acceleration of discovery-to-approval timelines by streamlining data-driven decision-making.
- Lower Development Costs: Fewer failed candidates and optimized trial designs reduce wasted resources.
- Improved Resource Utilization: Better prediction of demand, recruitment, and manufacturing capacity leads to leaner operations.
- Faster Time-to-Market: Rapid translation of scientific breakthroughs into marketed therapies improves patient access and industry competitiveness.

AI in Drug Discovery and Screening

Artificial intelligence (AI) is becoming a vital component of contemporary drug discovery, significantly reducing the reliance on trial-and-error methods and enabling a data-driven approach to finding new therapeutics. By integrating diverse biological, chemical, and clinical datasets, AI facilitates target identification, compound screening, hit optimization, and the discovery of novel chemical entities. Below is a detailed look at its major applications[32].

1. Target Identification and Validation

Target identification remains the critical first step in drug discovery, as it determines the biological macromolecule whose modulation is expected to produce therapeutic benefit. AI enhances this step by mining high-dimensional, heterogeneous datasets:

- Multi-Omics Integration: AI algorithms synthesize data from genomics, transcriptomics, proteomics, and metabolomics to identify disease-associated genes and pathways. For example, graph neural networks (GNNs) and knowledge graph-based models link gene expression profiles, protein–protein interaction networks, and disease ontologies to generate disease–gene association scores, revealing non-obvious therapeutic targets[33].
- Functional Genomics & CRISPR Screens: AI models analyze high-throughput CRISPR-Cas9 loss-of-function or gain-of-function screens, ranking candidate genes by their predicted therapeutic impact[34].
- Causal Inference & Systems Biology: ML-based causal inference models determine whether a target is likely to be a disease driver versus a downstream biomarker, increasing the probability of clinical success [35].
- Example: BenevolentAI's knowledge graph uses natural language processing (NLP) to ingest millions of scientific publications and patents, mapping drug–target–disease relationships that have led to the identification of new targets in atopic dermatitis and ALS [36].

2. Virtual Screening and Docking

AI-driven virtual screening dramatically reduces the number of compounds requiring physical synthesis and biological testing:

- AI-Enhanced Docking: Deep learning-based scoring functions (e.g., DeepDock, GraphDTA) learn from known protein–ligand complexes to predict more precisely determining binding affinities than traditional force-field or physics-based scoring functions.
- Ultra-Large Library Screening: ML models pre-filter billions of molecules from virtual libraries, eliminating molecules unlikely to bind or with undesirable ADMET profiles before docking[37].
- Pose Prediction: AI models outperform classical docking tools in predicting ligand binding poses, especially for targets with flexible binding sites.
- Case Example: Atomwise's AI platform has been used to screen >16 billion compounds for Ebola and multiple oncology targets, yielding hit rates several-fold higher than random screening[38].

3. De Novo Drug Design

The use of generative AI (8,14,15) models for drug design is among the top revolutionary advances:

- **Generative Models:** Variational autoencoders (VAEs), generative adversarial networks (GANs), and transformer-based models prefered**novelmolecules**designing as in reference to physicochemical and biological properties.
- **Multi-Objective Reinforcement Learning:** These models optimize for multiple parameters—potency,

selectivity, solubility, permeability, metabolic stability, and synthetic accessibility—simultaneously.

- **Synthetic Route Prediction:** AI suggests retrosynthetic routes for proposed molecules, ensuring **synthetic feasibility** and reducing development cycles.
- Example: Insilico Medicine (8,14)'s Chemistry42 platform generated INS018_055, the first fully AI-designed drug to enter Idiopathic pulmonary fibrosis phase II clinical studies[39].

4. Phenotypic Screening

Phenotypic screening focuses on identifying compounds that produce desirable cellular or organismal responses, even when the target is unknown.

- **AI-Powered Image Analysis:** Convolutional neural networks (CNNs) analyze**high-content microscopy images**, classifying subtle cellular morphologies and alterations in phenotype that are invisible to human observers.
- **Unbiased Mechanism of Action (MoA) Discovery:** Through clustering high-dimensional phenotypic profiles, artificial intelligence (AI) enables the ability to understand drug processes and deduce MoA or uncover biological pathways unknown before. This feature allows for the creation of first-in-class drugs.
- Multi-Modal Data Integration: AI-driven algorithms score candidate drugs with disease-relevant molecular signatures and make mechanistic predictions by integrating transcriptomic and proteomic data with phenotypic readouts.
- Applied Example: Recursion Pharmaceuticals generates robust "cellular response maps" by integrating AI-powered image analytics with high-throughput automated microscopy. Across numerous disease areas, such as oncology, rare genetic diseases (10,11), and inflammatory conditions, such platforms enable the rapid identification of new therapeutic avenues[40].

Strategic Advantages

Artificial intelligence (AI) methods have a synergistic impact on the pharmaceutical discovery pipeline when utilized end-to-end. They enhance success rates by prioritizing genetically and mechanistically validated targets, lower screening costs by focusing on high-value candidate compounds, and accelerate hit identification by compressing timelines from years to months. In addition to traversing the limits of current compound collections, AI also facilitates easier exploration of previously unvisited chemical space. Drug discovery is consequently being reengineered from a largely empirical, time-consuming process into a rational, hypothesis-based, and iterative process by the integration of AI, large biomedical datasets, and experimental biology[41].

Prediction of Bioactivity and Toxicity

Modern drug discovery increasingly depends on in silico prediction of both a compound's desired bioactivityand potential toxicity. Accurate prediction early in the pipeline helps reduce costs, avoid failures in later stages, and aligns with ethical and regulatory trends toward minimizing animal usage[42].

1. Bioactivity Prediction

- **Molecular Graph Representations & Deep Learning:**Graph neural networks (GNNs) and related deep learning architectures represent molecules as graphs (atoms = nodes, bonds = edges), enabling the model to learn both atomic and relational features. This is often superior to traditional QSAR models that depend on hand-crafted descriptors. For example, GraphEGFR integrates molecular graphs + fingerprint-based features and employs multi-task and transfer learning to improve prediction of inhibitor bioactivityagainst both wild-type and mutant EGFR family proteins, especially under data scarcity[43].
- **Transfer Learning & Pre-training:**Pre-training on large general chemical/bioactivitydatasets (e.g. ChEMBL) and then fine-tuning for specific targets helps improve generalization, especially when target-specific data are limited. One example is MolPMoFiT, which pre-trains a molecular structure prediction model and then fine-tunes for various QSAR/QSPR tasks; it showed strong performance without needing endpoint-specific hyperparameter tuning[44].
- **Uncertainty Quantification and Active Learning:** Some models adopt Bayesian or semi-supervised learning to also assess prediction uncertainty (i.e., how confident the model is). This helps in prioritizing compounds and avoiding overtrusting predictions in unexplored chemical space[45].
- **Benchmarking & Models Without Descriptors:** There are works demonstrating that GNNs or graph convolutional neural networks can perform QSAR tasks (e.g. mutagenicity prediction) without explicit molecular descriptors, by learning directly from graph structure (or SMILES representations converted into graphs). These approaches sometimes include attention mechanisms and uncertainty estimation[46].

2. Toxicity Prediction

- **ADME / Organ/Tissue-Specific Toxicities**.: AI models now predict liabilities like hepatotoxicity, immunotoxicity, cardiotoxicity(especially via hERG channel inhibition), mutagenicity, and others. For example, a multi-model deep learning platform was used to assess carcinogenicity, hERG-mediated cardiotoxicity, mutagenicity, and hepatotoxicity; models had AUCs in the range ~0.69 to ~0.88 depending on endpoint [47].

- **Cardiotoxicity&hERG Blocking.:**The CardioTox net model is an ensemble deep-learning framework that classifies compounds as hERG blockers or non-blockers, combining different base models and features such as fingerprints, SMILES embeddings, etc., to achieve improved performance. Also, recent studies (e.g., "hERG-toxicityprediction using traditional ML and advanced DL techniques") show that GNN-based models outperform classical ML approaches for large datasets of hERG channel data[48].

- **Early-Stage Filtering.**

  In an example: for artemisinin derivatives, a scheme combining cardiotoxicity, hepatotoxicity, immunotoxicity, tumorigenicity, mutagenicity, and reproductive toxicityfilters was used; 7 out of 374 compounds passed the filtering and showed very low hERG interaction in vitro[49].

- **Regulatory / Ethical Trends: New Approach Methodologies (NAMs).:**The U.S. FDA has announced a roadmap to reduce, refine, or replace animal testing using NAMs, including AI-based computational models, organoids, organ-on-chips, etc., especially in preclinical safety/toxicitystudies. These efforts are intended to allow submission of non-animal safety data in investigational new drug (IND) applications, where appropriate[50].

Strategic Impacts & Challenges

- **Trade-off among accuracy, coverage, and interpretability.**Superior accuracy models may not always be interpretable; uncertainty estimation helps, but interpretability remains key especially for regulatory acceptance[51].

- **Domain of applicability.** Predictive models tend to perform worse when applied to chemical space dissimilar from their training data; transfer learning and multi-task models help ameliorate this[52].

- **Data quality and bias.** Public datasets often have imbalanced labels or limited chemical diversity; high-quality curated data remains essential[53].

- **Regulatory acceptance.**The NAMs program, Togetherled the charge by the FDA and NIH, is an example of a paradigm shift in the direction of the combination of computer modeling with human biology–based in vitro systems. The successful execution of these techniquesdepends on onstrict validation and defining standardized protocols to facilitate consistency, reliability, and regulatory approval[54].

AI in Clinical Trials and Product Management

Artificial intelligence (AI) is contributing to modern drug development, impacting not just the design and operational management of clinical trials but also post-approval surveillance of therapeutic modifications. By deploying advanced computational algorithms to large amounts of real-world evidence and varied unstructured data streams, AI facilitates enhanced trial execution efficiency, reduces overall development costs, improves regulatory and clinical decision-making, and accelerates the production of actionable insights across the pharmaceutical lifecycle[55]

Clinical Trials

Patient Stratification and Matching

- Precision subpopulation Classification: In order to identify which patient populations are most likely to respond to treatment, artificial intelligence (AI) can aggregate and analyze genetic, phenotypic, and clinical information. Precision classification minimizes clinical outcome variation and maximizes enrichment of trial populations with probable responders[56].
- Optimization of clinical site selection: Machine learning algorithms can determine site performance, patient population density, logistics, and investigator networks using past trial data. They can then recommend trial sites with high recruitment potential and low participant attrition risk[57].

Synthetic Control Arms & Digital Twins

- **Synthetic Control Arms (SCAs):** SCAs are control groups that are generated using real-world evidence (RWD) or historical clinical trial data as compared to enrolling new placebo or control patients. They are especially convenient in scenarios where the employment of placebo arms is ethically questionable or in orphan diseases. Interestingly, the FDA cleared Medidata's Acorn AI SCA for a hybrid external control design, and it has been applied in several studies, such as a Phase III registrational trial of recurrent glioblastoma (rGBM)[58].

- VEIL and Bayer FCT as an Applied Example. AI: In collaboration with VEIL.AI through its Future Clinical Trials (FCT) program, Bayer used clinical trial data as well as anonymised registry data to develop SCAs. This approach was meant to reduce the burden on participants while speeding up the execution of the trial[59].

- **Digital Twin cohorts.** In one proof-of-concept study for chronic graft-versus-host disease (cGvHD), AI-derived digital twin cohorts replaced or augmented standard of care control arms. The platform built digital patient profiles from very large data sources to mirror what control group outcomes would be[60].

3. Dropout Prediction

- **Predicting patient dropout risk.**Implementation of employed machine learning models for various therapeutic contexts as a resultant for the patience who dropped out, allowing remedial interventions. For example, in substance use disorder treatment, the AI models which have been used for social platform + clinical intake data predicted 90-day dropout with AUC ~0.81, outperforming traditional psychometric scales[61].

- **Addiction recovery programs.** In the Houston "HEROES" opioid use disorder program (700+ participants), ML models predicted dropout at 90 and 120 days with sensitivity ~0.81/0.86 and specificity ~0.65/0.66 respectively, enabling targeted retention strategies[62].

Product Management & Post-Market Surveillance

Artificial intelligence (AI) is increasingly transforming the design and optimization of pharmaceutical excipients, particularly in the development of co-processed excipients (CPEs), by enabling data-driven formulation strategies and predictive material science. Machine learning (ML) and deep learning (DL) algorithms analyze large datasets comprising physicochemical properties, compressibility indices, flow characteristics, hygroscopicity, particle morphology, and compatibility profiles to predict optimal excipient combinations for desired tablet performance. In co-processed systemswhere two or more excipients are engineered at the sub-particle level to achieve synergistic functionality, AI facilitates multi-objective optimization, balancing parameters such as tensile strength, disintegration time, dissolution rate, and stability [63]. Predictive modeling tools can forecast excipient–drug interactions, moisture sensitivity, and compaction behavior before physical trials, thereby reducing experimental workload and development time. Furthermore, AI-assisted design of experiments (DoE), material informatics, and process analytical technology (PAT) integration enable real-time monitoring and quality-by-design (QbD) compliance in manufacturing. As generative AI expands into material design, it may soon assist in conceptualizing next-generation smart co-processed excipients with tailored release profiles and enhanced manufacturability, marking a significant evolution in pharmaceutical formulation science [64].

1. Market Forecasting and Competitor Analysis

- **Demand forecasting.** AI methods analyze historical sales, epidemiological trends, payer coverage policies, and demographic shifts to predict future demand for novel therapies. These forecasts inform manufacturing scale, supply chain planning, and commercial rollout strategies.

- **Competitive intelligence.** NLP and text mining of patent filings, clinicaltrials.gov, publications, conference proceedings, and regulatory filings allow early detection of competitor pipeline moves, MOAs under investigation, or changes in standards of care.

2. Pharmacovigilance & Safety Signal Detection

Artificial intelligence (AI) is enhancing pharmacovigilance (PV) by improving the detection and management of adverse drug reactions (ADRs). Machine learning (ML) and natural language processing (NLP) analyze data from electronic health records, safety databases, and social media to identify safety signals earlier than traditional methods. AI also supports automated case processing, signal detection, causality assessment, and risk prediction, reducing manual workload and improving post-marketing drug safety monitoring [65].

Artificial intelligence (AI) is increasingly being applied in **materiovigilance**, the systematic monitoring of adverse events associated with medical devices. AI techniques such as machine learning (ML) and natural language processing (NLP) analyze large volumes of data from device registries, electronic health records (EHRs), complaint databases, and post-marketing surveillance reports to detect early safety signals. Unlike traditional passive reporting systems, AI can identify hidden patterns, predict device malfunction risks, and classify incident severity automatically. It also assists in real-time signal detection, trend analysis, root-cause identification, and automated case triaging, improving regulatory compliance and response time. By integrating real-world evidence (RWE) and predictive analytics, AI enhances proactive risk management, ultimately improving patient safety and device performance monitoring [66].

Adverse Drug Reaction (ADR) prediction. ML models applied to hospitalization data, EHRs, lab values, and comorbidity information can predict ADR risks. A systematic review found many such models, but also noted limitations in autitsand generalizability.

Unstructured data mining. Use of NLP on social programming, patient forums, biomedical literature, case reports to detect safety signals. For example, an AI approach applied to patient reviews in the *Levothyrox* case (France) used social media forum data + sentiment analysis + NLP + CNN models to detect increases in ADR reports after formula changes. Model accuracy ~75%, Databases & package tools. Enhancement of novel tools such as MDDC (Modified Detecting Deviating Cells) in R/Python for adverse event identification in large pharmacovigilancedatabases like FAERS, VigiBase, etc[67].

3. Strategic Benefits & Challenges

- **Benefits**: Reduced trial durations, fewer sample population needed for controlled better patient retention, more efficient allocation of marketing resources, earlier detection of adverse events and product risks, improved trust with regulators and payers[68].

- **Challenges**: Access to high-quality real-world data; privacy, anonymization, and regulatory compliance; bias / representativeness of datasets; explainability of AI models; regulatory acceptance and guidelines for synthetic arms, digital twins, and AI-derived evidence[69-70].

Challenges and Future Perspectives

Despite rapid advances and promising applications, Artificial Intelligence (AI) in drug discovery confronts significant obstacles. Addressing these challenges, and pursuing future directions that resolve or bypass them, will be key to realizing AI's full transformative potential[71].

Key Challenges

Data Quality & Bias

- Many datasets used for training AI models are incomplete, non-standardized, or skewed toward well-studied targets, populations, or disease types. Under-representation of rare diseasesor minority populations leads to models that generalize poorly[72].
- Label noise, missing values, inconsistent assay protocols, batch effects, and bias introduced during data collection (such as over-sampling certain chemical scaffolds) further degrade model reliability and may produce spurious associations[73].

Interpretability and Transparency

- Deep learning algorithms are frequently "black boxes," providing little information about how predictions are made. This lack of transparency hinders **regulatory approval**, clinician or stakeholder trust, and error diagnosis.
- Explainable AI (XAI) methods (feature attribution, attention maps, counterfactuals etc.) are being developed, but often they trade off predictive performance, or are difficult to scale to multi-modal (e.g., chemical + biological + clinical) data[74].

Benchmarking & Reproducibility

- Lack of standardized benchmarks: many studies use proprietary or idiosyncratic datasets, making cross-comparison difficult. The evaluation metrics may vary (AUC, accuracy, F1, etc.), with differing thresholds, which complicates interpretation of model performance.
- Reproducibility issues: differences in preprocessing, hyperparameter settings, random seeds, hardware (GPUs, TPUs), and code availability lead to results that cannot be replicated. These challenges are especially acute when models are evaluated only on retrospective datasets.

Ethics & Privacy

- Protected health information (PHI) is sensitive; ensuring consent, anonymization, and data security is essential. Data sharing is often limited by privacy laws, institutional restrictions, and ethical concerns.
- Algorithmic bias may exacerbate health disparities: if certain demographic or genetic groups are underrepresented, AI predictions may perform worse for them, raising fairness issues[75].

Integration, Infrastructure & Regulatory Hurdles

- Incorporating AI tools into established R&D pipelines requires changes in infrastructure, data governance, staff training, and sometimes culture. Many pharmaceutical organizations are siloed, which limits data flow and cross-functional cooperation.
- Regulatory frameworks for AI-derived predictions, synthetic control arms virtual twin models, etc., are still evolving. Clear guidelines, standards of evidence, validation, auditing, and explainability will be needed[76].

Future Perspectives and Research Directions

Hybrid Models Combining Mechanistic Simulation with ML

- Models that couple **mechanistic or physics-based simulation** (e.g., quantitative systems pharmacology, PBPK, molecular dynamics) with data-driven ML are promising. They can leverage prior knowledge, reduce overfitting, and increase interpretability. This hybrid approach is specially done for some of diseases and early-discovery where data are sparse.
- Examples: "Programmable Virtual Humans", a perspective work, advocates multiscale models combining molecular to phenotypic level simulations augmented with AI to bridge translational gaps[77].

Transparent Validation Pipelines and Standardization

- Adoption of standardized datasets, open benchmarks, shared preprocessing pipelines, and mandatory reporting of hyperparameters, random seeds, performance metrics (with confidence intervals) will improve reproducibility and comparability.
- Public platforms (e.g. PapersWithCode), open-source code and model cards that document data provenance, intended domain, limitations, biases. Regulatory bodies could adopt guidelines requiring AI model documentation (akin to Good Laboratory Practice or clinical trial reporting standards)[78].

Multi-Modal and Foundation Models Capable of Cross-Domain Reasoning

Development of foundation models trained across chemistry, biology, clinical data, imaging, etc., which can be fine-tuned for specific tasks, facilitating transfer learning, zero-shot or few-shot learning in under-studied areas. These models can help integrate and reason across modalities and scales. For some, models that integrated models use single-cell / spatial omics with imaging and clinical endpoints[79].

Ethical, Fairness, and Privacy Enhancements

- Method such as **federated learning**, **differential privacy**, chemically generated data, and robust anonymization will be critical to allow sharing of sensitive data across institutions and geographies without compromising privacy.

- Fairness audits and bias detection pipelines must be embedded in model development to ensure equitable performance across demographic, genetic, and geographical groups[80].

Regulatory & Policy Evolution

- Regulatory agencies (FDA, EMA, etc.) need to provide clearer guidelines for AI tools in drug discovery—regarding what validation, transparency, human interpretability, and accountability are required.
- Encouragement of prospective validation studies (not just retrospective), public-private partnerships to generate high-quality diverse datasets, and incentives for sharing data, especially in less served disease areas[81].

CONCLUSION

AI is reshaping every stage of the drug discovery and development pipeline. By accelerating early discovery, improving predictive modeling, and optimizing clinical development, AI has the capability to deliver convineant and more effective drugs faster than ever before. Ongoing cooperation among academic institutions, industry stakeholders, and regulatory bodies will be crucial to fully achieve this potential.

REFERENCES

1. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–589.
2. Abramson J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024.
3. Pandey M, et al. The transformational role of GPU computing and deep learning in drug discovery. Nat Mach Intell. 2022;4:123–135.
4. Blanco-González A, et al. The role of artificial intelligence in drug discovery: challenges and opportunities. Front Pharmacol. 2023;14:1225120.
5. Hasselgren C, et al. Artificial intelligence for drug discovery: are we there yet? Annu Rev Pharmacol Toxicol. 2024;64:21–43.
6. Kim H, et al. Artificial intelligence in drug discovery: a comprehensive review. Pharmaceuticals (Basel). 2021;14:1113.
7. Cavasotto CN, et al. Machine learning toxicity prediction: latest advances and applications. Drug Discov Today. 2022;27:123–135.
8. Ferreira FJN, et al. AI-driven drug discovery: a comprehensive review. Expert Opin Drug Discov. 2025;20:45–62.
9. Guo W, et al. Machine learning and deep learning models for toxicity prediction. Front Pharmacol. 2023;14:1182456.
10. Lu X, et al. Artificial intelligence for optimizing recruitment and conduct of clinical trials. Front Med. 2024;11:1302156.
11. Ismail A, et al. The role of artificial intelligence in accelerating patient recruitment in clinical trials. Contemp Clin Trials. 2023;128:107164.
12. Qian L, et al. Enhancing clinical trial outcome prediction using machine learning. Clin Pharmacol Ther. 2025;117:345–357.
13. Paul SM, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov. 2010;9:203–214.
14. Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018;17:97–113.
15. Zhavoronkov A, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol. 2019;37:1038–1040.
16. Ekins S, et al. Exploiting machine learning for end-to-end drug discovery and development. Nat Mater. 2019;18:435–441.
17. Vamathevan J, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18:463–477.
18. Stokes JM, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180:688–702.
19. Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects. Drug Discov Today. 2019;24:773–780.
20. Exscientia. AI-designed drug discovery platform overview. Drug Discov Today. 2022;27:100–110.
21. Walters WP, Murcko MA. Assessing the impact of generative AI on medicinal chemistry. J Med Chem. 2020;63:865–878.
22. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018;23:1241–1250.
23. Elton DC, et al. Deep learning for molecular design—a review of generative models. Mol Syst Des Eng. 2019;4:828–849.
24. Kirchmair J, et al. Predicting drug metabolism: experiment and/or computation? Nat Rev Drug Discov. 2015;14:387–404.
25. Rajkomar A, et al. Machine learning in medicine. N Engl J Med. 2019;380:1347–1358.
26. Berry DA. Bayesian adaptive clinical trials. Nat Rev Drug Discov. 2011;10:27–36.
27. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25:44–56.
28. Harpaz R, et al. Data mining for adverse drug events. Clin Pharmacol Ther. 2012;91:1010–1021.
29. Ierapetritou MG, et al. Artificial intelligence in pharmaceutical manufacturing. AIChE J. 2020;66:e16861.
30. Bate A, Edwards IR. Data mining techniques in pharmacovigilance. Drug Saf. 2006;29:393–397.
31. Krittanawong C, et al. Deep learning for cardiovascular medicine. Eur Heart J. 2017;38:2059–2067.
32. Gawehn E, et al. Deep learning in drug discovery. Mol Inform. 2016;35:3–14.
33. Zitnik M, et al. Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics. 2018;34:i457–i466.
34. Meyers RM, et al. Computational correction of copy number effect improves CRISPR screen specificity. Nat Genet. 2017;49:1779–1784.
35. Pearl J. Causal inference in statistics: an overview. Stat Surv. 2009;3:96–146.
36. BenevolentAI. Knowledge graph applications in drug discovery. Drug Discov Today. 2021;26:200–208.
37. Gorgulla C, et al. An open-source drug discovery platform enables ultra-large virtual screens. Nature. 2020;580:663–668.
38. Atomwise Inc. AI-driven virtual screening advances. J Chem Inf Model. 2021;61:512–520.
39. Zhavoronkov A. Artificial intelligence for drug discovery, biomarker development, and generation of novel chemistry. Mol Pharm. 2018;15:4311–4313.
40. Recursion Pharmaceuticals. AI-enabled phenotypic drug discovery. Nat Biotechnol. 2022;40:121–130.
41. Walters WP. Virtual chemical libraries and AI. J Med Chem. 2019;62:1116–1124.

42. Ekins S. The next era: deep learning in pharmaceutical research. Pharm Res. 2016;33:259–262.
43. Jiang D, et al. Could graph neural networks learn better molecular representation? J Chem Inf Model. 2021;61:2730–2740.
44. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9:1735–1780.
45. Gal Y, Ghahramani Z. Dropout as Bayesian approximation. ICML. 2016.
46. Duvenaud D, et al. Convolutional networks on graphs for learning molecular fingerprints. NIPS. 2015.
47. Wu Z, et al. MoleculeNet: a benchmark for molecular machine learning. Chem Sci. 2018;9:513–530.
48. Sun H, et al. Deep learning-based prediction of hERG blockade. Chem Res Toxicol. 2017;30:1624–1634.
49. Li X, et al. AI-based toxicity screening of artemisinin derivatives. Front Pharmacol. 2022;13:1002456.
50. US FDA. Advancing new approach methodologies (NAMs) roadmap. 2023.
51. Ribeiro MT, et al. Why should I trust you? Explaining classifier predictions. KDD. 2016.
52. Sheridan RP. Applicability domain in QSAR modeling. J Chem Inf Model. 2013;53:283–296.
53. Wallach I, et al. Most ligand-based classification benchmarks reward memorization. J Chem Inf Model. 2018;58:916–932.
54. NIH-FDA. Modernizing toxicity testing. Science. 2022;377:990–992.
55. Harrer S, et al. Artificial intelligence for clinical trial design. Trends Pharmacol Sci. 2019;40:577–591.
56. Miotto R, et al. Deep learning for healthcare. Brief Bioinform. 2018;19:1236–1246.
57. He Z, et al. Clinical trial site selection using AI. J Biomed Inform. 2020;108:103487.
58. Franklin JM, et al. Using real-world data to create synthetic control arms. Clin Pharmacol Ther. 2021;109:118–125.
59. Eichler HG, et al. External control arms in clinical trials. Lancet. 2020;396:1346–1348.
60. Liu X, et al. Digital twin modeling in medicine. NPJ Digit Med. 2022;5:42.
61. Che Z, et al. Predicting patient dropout using machine learning. J Biomed Inform. 2018;85:70–78.
62. Acion L, et al. Machine learning in addiction recovery programs. J Subst Abuse Treat. 2017;80:59–66.
63. Raval AM, Bhavsar PR, Pandya FU, Patel DK. Co-processed excipients in pharmaceutical formulation: advances, characterization, and applications. Asian J Pharm Res Dev. 2026;14(1):105–113. doi:10.22270/ajprd.v14i1.1705.
64. Raval AM, Suthar AM, Durani B, Thakar NJ, Zankhwala FM, Kushkiwala AM, et al. Smart co-processed excipient platforms: A novel strategy for multifunctional optimization of ibuprofen tablet formulations. J Appl Bioanal. 2025;11(15 Suppl):103–128. doi:10.53555/jab.v11si15.2109.
65. Mevada J, Patel K, Raval A. Role of pharmacovigilance in drug safety monitoring. World J Pharm Med Res. 2025;11(11):235–240. doi:10.5281/zenodo.17482904
66. Mevada J, Patel K, Raval A. Materiovigilance: from device failure to safety reform, the growing importance of materiovigilance systems. Int J Pharm Sci. 2026;4(1):962–974. doi:10.5281/zenodo.18207625
67. Norén GN, et al. Data mining for signal detection in pharmacovigilance. Pharmacoepidemiol Drug Saf. 2008;17:133–140.
68. Falwariya R, Jethva T, Raval AM, Lokhande D. Comprehensive review: microneedle patches – a painless revolution in transdermal drug delivery. *World J Pharm Med Res.* 2026;12(1):199–207. doi:10.5281/zenodo.18085578
69. Topol EJ. Preparing the healthcare workforce for AI. Nat Med. 2019;25:44–56.
70. Wiens J, et al. Do no harm: AI for healthcare challenges. Nat Med. 2019;25:1337–1340.
71. Ching T, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface. 2018;15:20170387.
72. Rathi S, Shah S, Raval AM, Patel D, Goswami A. Physicochemical characterization and in-vitro dissolution enhancement of ranolazine using solid dispersion method. J Emerg Technol Innov Res. 2019;6(3):866–883.
73. Bender EM, et al. On the dangers of stochastic parrots. FAccT. 2021.
74. Samek W, et al. Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. IEEE Signal Process Mag. 2017;34:87–95.
75. Zankhwala FM, Raval AM, Kushkiwala AM, Sarvaiya SP, Raval KK, Thakar NJ, et al. Formulation and evaluation of optimized polymer blends for diclofenac diethylamine transdermal system. Rev Diabet Stud. 2025;21(S9):701–708.
76. EMA. Guideline on computerised systems and electronic data in clinical trials. 2023.
77. Viceconti M, et al. In silico clinical trials: how computer simulation will transform medicine. Sci Transl Med. 2016;8:342ps13.
78. Mitchell M, et al. Model cards for model reporting. FAT*. 2019.
79. Moor M, et al. Foundation models for biomedical AI. Nat Med. 2023;29:1318–1328.
80. Kairouz P, et al. Advances in federated learning. Found Trends Mach Learn. 2021;14:1–210.
81. US FDA. Artificial intelligence/machine learning (AI/ML)-based software as a medical device action plan. 2021.