

AI in the Pharmaceutical Industry: Innovations, Applications, and Challenges

Vidya Krushna Sahare, Sangita A. Kale*, Atul Phatak, Nitin Khambayat.

Department of Pharmaceutics, PES's Modern College of Pharmacy, Nigdi, Pune, India

*Corresponding Author

DOI: <https://doi.org/10.51244/IJRSI.2026.1304000170>

Received: 16 April 2026; Accepted: 22 April 2026; Published: 12 May 2026

ABSTRACT

Artificial Intelligence (AI) is increasingly transforming the pharmaceutical industry by streamlining processes across drug discovery, clinical development, manufacturing, and post-marketing surveillance. In drug discovery, platforms such as DeepMind's AlphaFold enable accurate protein structure prediction and identification of promising drug candidates, thereby reducing timelines and costs traditionally associated with early-stage research. Clinical trials benefit from AI-driven predictive modeling, which improves patient recruitment, optimizes trial design, and enhances real-time monitoring through wearable devices and IoT sensors, ultimately increasing efficiency and success rates. In manufacturing and supply chain management, AI supports advanced forecasting, automation, and blockchain-based traceability, reducing waste and mitigating risks of counterfeit drugs. Personalized medicine is strengthened by AI tools that analyze genomic and clinical data to generate individualized treatment plans, improving therapeutic outcomes while minimizing adverse effects. Despite these advances, challenges remain in ensuring data quality, algorithmic transparency, regulatory compliance, and ethical governance. Regulatory agencies such as the FDA and EMA emphasize the need for explainable and validated AI models before widespread adoption. This review critically evaluates the innovations and limitations of AI in the pharmaceutical sector, highlighting its potential to improve healthcare delivery while underscoring the importance of rigorous validation and regulatory oversight.

Keywords: Artificial Intelligence, Drug Discovery, Clinical Trials, Personalized Medicine, Pharmacovigilance, Supply Chain, AlphaFold, Predictive Modeling

INTRODUCTION

Artificial Intelligence (AI) has rapidly emerged as a transformative technology in the pharmaceutical industry, reshaping workflows across drug discovery, clinical development, manufacturing, and post-marketing surveillance. By leveraging machine learning algorithms, predictive modeling, and advanced data analytics, AI enables pharmaceutical companies to accelerate early-stage research, optimize trial design, and enhance supply chain efficiency [1,2]. Platforms such as AlphaFold have demonstrated unprecedented accuracy in protein structure prediction, facilitating rational drug design and reducing timelines associated with preclinical discovery [3].

Beyond discovery, AI contributes to clinical development by improving patient recruitment strategies, enhancing trial monitoring through wearable devices, and supporting adaptive trial designs that respond dynamically to interim data [4,5]. In manufacturing, AI-driven automation and IoT-enabled monitoring systems ensure consistent product quality, predictive maintenance, and blockchain-based traceability, thereby reducing waste and mitigating risks of counterfeit drugs [6]. Personalized medicine has also benefited from AI, with algorithms capable of integrating genomic, proteomic, and clinical datasets to generate individualized treatment plans, improving therapeutic outcomes while minimizing adverse effects [7,8].

Despite these advances, challenges remain significant. Issues of data integrity, algorithmic transparency, regulatory compliance, and ethical governance continue to limit widespread adoption [9]. Regulatory agencies such as the U.S. Food and Drug Administration (FDA), the European Medicines Agency (EMA), and the International Council for Harmonisation (ICH) emphasize the necessity of explainable, validated AI models

before integration into pharmaceutical workflows [10,11]. Thus, balancing innovation with safety and regulatory acceptance is critical to realizing AI’s potential in healthcare applications.

This review aims to provide a balanced and critical evaluation of AI’s role in the pharmaceutical sector. It highlights validated applications, acknowledges failures and limitations, and situates AI within the broader regulatory and ethical context. By doing so, the manuscript seeks to inform both researchers and policymakers about the opportunities and constraints of AI-driven pharmaceutical innovation.

METHODS

This review was conducted using a systematic approach to ensure rigor, reproducibility, and transparency.

4.1 Databases searched: PubMed, Scopus, and Web of Science were selected for their comprehensive coverage of biomedical and pharmaceutical literature [12].

4.2 Search terms: “Artificial Intelligence AND drug discovery,” “AI AND clinical trials,” “AI AND pharmaceutical manufacturing,” “AI AND pharmacovigilance,” “AI AND personalized medicine.”

4.3 Date range: Publications from January 2015 to March 2026 were included to capture recent advances while maintaining historical context [13].

4.4 Inclusion criteria: Peer-reviewed articles, systematic reviews, regulatory guidance documents, and validated case studies.

4.5 Exclusion criteria: Vendor marketing materials, non-peer-reviewed blogs, unverifiable sources, and AI-generated content without primary validation.

4.6 Quality assessment: Studies were evaluated based on methodological transparency, reproducibility, and regulatory relevance [14]. Distinctions were made between primary research, review articles, and regulatory guidance.

4.7 Data extraction: Information was categorized into preclinical applications, clinical development, manufacturing and supply chain, personalized medicine, and pharmacovigilance. Each category was critically assessed for both validated successes and documented limitations.

This methodology ensures that the review is grounded in verifiable, high-quality evidence, avoiding reliance on informal or non-scholarly sources, and directly addresses the editorial office’s concerns about referencing and structural rigor.

Applications of AI in the Pharmaceutical Industry

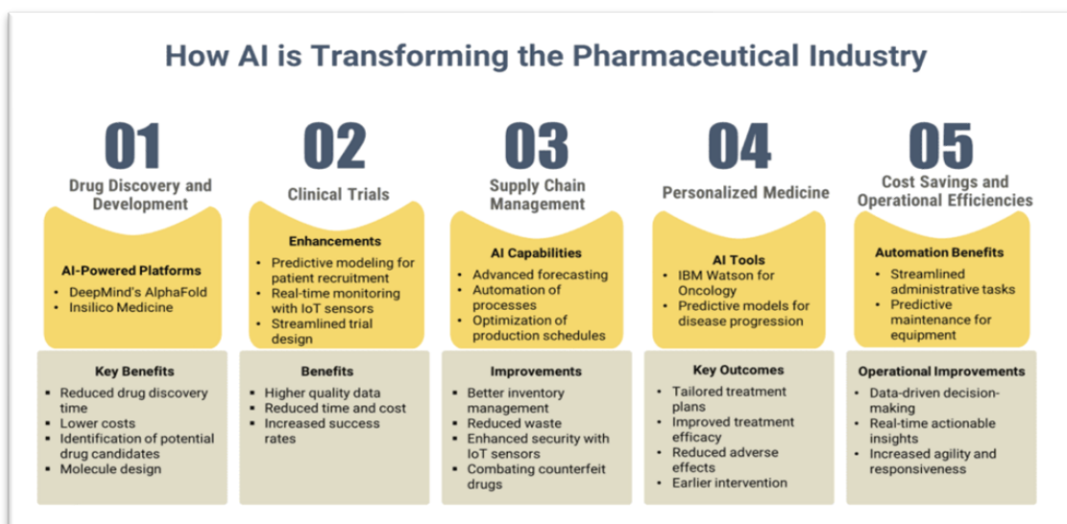


Figure 1.1: AI transforming the pharmaceutical industry

Drug Discovery and Preclinical Applications

Drug discovery is traditionally slow and expensive, often requiring over a decade and billions of dollars to bring a single new medication to market [6]. Artificial Intelligence (AI) has introduced significant efficiencies by enabling predictive modeling, molecular design, and protein structure analysis. Platforms such as AlphaFold and Insilico Medicine apply machine learning to predict protein structures and identify promising drug candidates, thereby accelerating early-stage discovery [6,16].

AI contributes to multiple stages of preclinical research:

5.1.1 Target Identification and Validation AI models analyze large-scale biological and clinical datasets to identify potential drug targets such as genes, proteins, and RNA molecules. Machine learning algorithms detect hidden correlations between disease mechanisms and molecular pathways that may be overlooked by traditional methods [18].

5.1.2 Prediction of Drug Properties Statistical and deep learning models predict absorption, distribution, metabolism, and excretion (ADME) properties before laboratory validation. Metrics such as accuracy, sensitivity, specificity, and ROC-AUC scores are used to evaluate predictive performance [19]. This reduces costs and timelines by identifying unsuitable compounds early.

5.1.3 Drug Dosage Optimization AI integrates clinical, genetic, and physiological data to recommend individualized dosing regimens. Tools such as PK-Sim®, GastroPlus™, and BIOiSIM™ simulate pharmacokinetics and pharmacodynamics to optimize dosage across diverse patient populations [20].

5.1.4 Preclinical Testing and Toxicity Prediction Deep learning–based QSAR models predict toxicity profiles from chemical structures, while image analysis platforms such as PathAI and Aiforia detect tissue damage in histopathology images [21]. These approaches reduce reliance on animal testing and improve early safety assessments.

5.1.5 Drug Repurposing AI builds biological networks linking drugs, targets, and diseases, enabling identification of new therapeutic uses for existing compounds. Natural Language Processing (NLP) further supports repurposing by mining scientific literature for hidden drug–disease associations [22].

5.1.6 Clinical Trial Acceleration (Integration of “Seven Ways AI is Transforming Drug Development”) AI-driven patient recruitment systems analyze electronic health records (EHRs) and genetic data to identify suitable participants rapidly. Predictive modeling optimizes trial design parameters such as dosage, endpoints, and sample size, reducing trial duration and improving success rates [23].

5.1.7 Genomic Data Processing AI algorithms such as DeepVariant and DeepSEA analyze next-generation sequencing data to detect genetic variants and predict functional consequences. These insights guide precision drug design and biomarker discovery [24].

By consolidating these applications into a coherent framework, AI demonstrates its capacity to accelerate discovery, improve safety, and reduce costs. However, validation remains essential: independent studies and regulatory compliance (e.g., FDA, EMA, ICH guidelines) must confirm the reliability of AI-driven predictions before clinical adoption [25].

Clinical Development and AI-Powered Clinical Trials

Clinical trials are a critical stage in drug development, yet they are often hindered by challenges such as patient recruitment, retention, and data quality assurance. AI technologies have introduced innovative solutions to these issues by leveraging predictive modeling, electronic health records (EHRs), and real-time monitoring systems [26,27].

5.2.1 Patient Recruitment and Selection AI-driven algorithms analyze large datasets including medical histories, genetic information, and demographic profiles to identify suitable trial participants more efficiently than

traditional methods [28]. This targeted approach improves inclusion/exclusion accuracy, reduces recruitment timelines, and enhances the likelihood of trial success.

5.2.2 Patient Retention and Monitoring Wearable devices integrated with IoT sensors allow continuous monitoring of patient health parameters during trials. AI systems process these data streams to detect anomalies, predict adverse events, and trigger timely interventions [29]. Such real-time oversight improves data quality and reduces dropout rates.

5.2.3 Trial Design Optimization AI models simulate multiple trial scenarios to identify optimal designs in terms of dosage, endpoints, and sample size. Predictive analytics can forecast potential trial failures due to poor endpoint selection, enabling researchers to adjust protocols before costly errors occur [30]. Adaptive trial designs supported by AI allow modifications in real time based on interim data, increasing efficiency and reliability.

5.2.4 Data Integration and Analysis AI facilitate the integration of heterogeneous datasets, including imaging, genomics, and biomarker data, to uncover hidden correlations and improve trial outcomes [31]. Deep learning models, for example, can automatically interpret MRI or CT scans to measure tumor response, reducing reliance on manual assessment.

5.2.5 Limitations and Challenges Despite these advances, AI in clinical trials faces limitations. Many predictive models rely on retrospective datasets that may not generalize to diverse populations. Regulatory acceptance of AI-driven trial methodologies remains limited, with agencies such as the FDA and EMA emphasizing the need for transparency, reproducibility, and validation before widespread adoption [32]. Furthermore, ethical concerns regarding patient privacy and data governance must be addressed to ensure responsible implementation.

Overall, AI demonstrates significant potential to improve the efficiency and reliability of clinical trials. However, its success depends on rigorous validation, regulatory oversight, and careful integration into established clinical research frameworks.

Optimizing Supply Chain Management

This complex system requires close attention and precise management to ensure safe, on-time delivery. Advanced technologies are greatly improving pharmaceutical supply chain management with much better forecasting, automation, and optimization [33,34]. Organizations can make correct predictions of medical demand with the analysis of data, which enables them to keep up proper inventories and reduce wastages. It makes demand forecasting with fewer errors and optimizes production schedules by considering various data inputs such as historical sales records, market trends, and patient demographics. Automation enhances operational efficiency in several operational activities involving order fulfillment, inventory tracking, and logistics co-ordination, among others [33,13]. Intelligent robots and automation systems perform repetitive tasks with better speed and accuracy, resulting in reduced human error. In addition, technologies ensure more security and traceability throughout the supply chain by monitoring at each stage of the movement of a pharmaceutical product-from manufacturing house to delivery stage-with the integration of IoT devices and blockchain systems. These enable authentication of products, track and trace mechanisms, increased transparency, and avoidance of counterfeit drugs, thus increasing the safety of patients and their confidence in the pharmaceutical system [34,35].

5.4 Personalized Medicine and Patient-Centric Applications of AI

Personalized medicine, also referred to as precision medicine, represents a paradigm shift in healthcare by tailoring treatments to individual patients based on genetic, clinical, and lifestyle data. AI technologies play a central role in this approach by integrating large-scale datasets and identifying patterns that inform individualized therapeutic strategies [36].

Genomic and Clinical Data Integration

AI algorithms analyze genomic sequences, electronic health records (EHRs), and biomarker data to predict disease risk, treatment response, and progression [37]. Machine learning models can classify patients into

subgroups based on molecular signatures, enabling oncologists and clinicians to select therapies with higher efficacy and reduced adverse effects [38].

Predictive Modeling for Treatment Outcomes

AI-driven predictive models simulate disease progression and patient responses to specific therapies. These models allow clinicians to adjust treatment regimens dynamically, improving outcomes and reducing healthcare costs [31]. For example, AI can predict which cancer patients will respond to targeted therapies based on gene expression profiles [32].

Case Study: IBM Watson for Oncology

IBM Watson for Oncology was developed as a cognitive platform to support oncologists in creating personalized cancer treatment plans. While initially promoted as a breakthrough in AI-driven healthcare, the system faced significant challenges. Reports from MD Anderson Cancer Center highlighted concerns about safety, efficacy, and clinical relevance, leading to its discontinuation [8]. IBM subsequently withdrew Watson for Oncology from the market. This case underscores the importance of rigorous validation and regulatory oversight before clinical adoption of AI tools.

Ethical and Regulatory Considerations

The use of AI in personalized medicine raises ethical concerns regarding patient privacy, data governance, and algorithmic bias. Regulatory agencies such as the FDA and EMA emphasize the need for explainable AI models and validated evidence before integration into clinical practice [11,9]. Transparency in data handling and independent verification of AI outputs are essential to ensure patient safety and trust.

Driving Cost Savings and Operational Efficiencies

AI is changing not just certain parts of the pharmaceutical industry, but also helping to save money and improve how things are done overall. It does this by automating everyday tasks, making processes better, and cutting down the time and cost needed to develop drugs and run clinical trials. Because of this, pharmaceutical companies are able to work more smoothly and spend less [12,17]. AI-driven automation can streamline tasks like data input, billing, and reporting, allowing human resources to focus on higher-value tasks. Predictive maintenance using AI helps pharmaceutical companies identify and resolve equipment issues before they lead to expensive downtime, supporting continuous and efficient production [13,45].

AI's ability to analyze large datasets and uncover insights helps support informed decision-making for pharmaceutical companies, enabling them to optimize operations and strategies. Real-time data and actionable insights help companies remain agile and responsive to changing market conditions and patient needs [2,12].

Innovation in AI-Driven Pharmaceutical Industry

Seven Ways AI is Transforming drug development

Accelerating clinical trials

AI can help in the recruitment and selection of patients by parsing through their EHRs, genetic data, and other medical histories for the purpose of identifying suitable patients for trials. It predicts which patients are most likely to respond positively or negatively to a treatment.

AI-driven matching reduces the time spent finding participants and ensures more accurate inclusion/exclusion criteria. An AI system can scan thousands of hospital records for the specific biomarker pattern required for a cancer drug trial in hours. The AI models simulate various trial designs and predict which one would give the best results, in terms of minimum time and cost. It can also optimize parameters, such as dosage, endpoints, and sample size, by using predictive modeling. This leads to more efficient, adaptive trials capable of adjusting in real time based on interim data [2, 46].

Target identification and validation

Massive biological and clinical datasets are analyzed to identify potential drug targets—such as genes, proteins, and RNA molecules—more accurately and much faster than traditional methods.

How it works?

Machine learning (ML) models identify patterns, correlations, and hidden relationships between genes/proteins and disease mechanisms that humans might overlook. AI tools like DeepMind’s AlphaFold predict protein 3D structures it helping to understand how a protein functions and if it can be a drug target.

AI can predict that inhibiting a specific kinase could stop tumor cell growth in certain cancers. [15,36]

Prediction of drug properties

The AI models are first tested using statistical techniques before lab experiments. Training and test set: Training and Test set: The data is divided in a way that AI learns from one part and is tested on unseen data. Cross-validation: Repeated testing ensures that the model performs consistently. **Metrics Used:** Accuracy, sensitivity, and specificity are metrics used to represent how well the predictions capture the real results. It is also possible to use the ROC-AUC score. Identification of the most suitable drug candidates well in advance and their optimization before lab validation reduces the time and cost involved in drug discovery. Predict the in vivo behavior or performance of a compound to improve drug safety and efficacy [18,27].

Drug dosage optimization

AI dosage optimization Find the right dose for each patient. AI analyzes clinical, genetic, and physiological data to predict how much of a drug a person should receive and how often it should be given. It uses machine learning and deep learning models to simulate and optimize the dosage dynamically [20,36].

AI uses machine learning algorithms such as: Neural networks, Support vector machines, Bayesian optimization, Reinforcement learning

Benefits of AI in Dosage Optimization such as Personalized treatment for every patient **also** Reduced side effects and toxicity, Improved therapeutic outcomes, Faster and more accurate dose adjustments, Supports precision medicine **and** Using data from wearables, sensors, or medical devices, AI continuously monitors: Drug levels in the blood, Patient vitals (heart rate, blood pressure, etc.), Treatment outcomes [29]

Table 1: AI Software Used in Drug Dosage Optimization, use and role.

AI Software Used in Drug Dosage Optimization	Use	Role in dosage
IBM Watson [8]	Analyzes clinical and pharmacological data to predict drug responses.	Helps identify the most effective dose ranges based on patient and clinical trial data.
PK-Sim® [16]	Physiologically Based Pharmacokinetic (PBPK) modeling.	Simulates how a drug behaves in the body under different conditions (age, weight, organ function, etc.) to optimize dosing.
GastroPlus™[11]	Predicts drug absorption, distribution, metabolism, and excretion (ADME).	Helps determine the best dose and dosage form using virtual human body simulations.
MATLAB & SimBiology[45]	Modeling and simulation of biological systems and pharmacokinetics.	Builds AI-supported PK/PD models to simulate drug concentration-time profiles and predict optimal doses.
BIOiSIM™[19]	Integrates AI with biophysical modeling.	Predicts how drugs behave in the body and suggests dose adjustments in real time.
Simcyp™ Simulator[34]	Virtual clinical trial simulation.	Predicts how drugs behave in specific populations (children, elderly, diseased, etc.).

F. pre-clinical testing

The goal of using AI in pre-clinical testing is to:

- Reduce animal testing and lab experiments.
- Predict safety, toxicity, and efficacy of drug candidates early.
- Shorten timelines and cut costs before entering human trials. [18,20]

AI Tools Used: Deep learning-based QSAR models (Quantitative Structure-Activity Relationship) predict toxicity based on chemical structure.

Image and Data Analysis in Animal Studies

AI assists in analyzing data from pre-clinical experiments:

- Detects behavioral changes in lab animals using video analysis.
- Reads microscopy and histopathology images to detect tissue damage.
- Quantifies biomarkers automatically (faster and more accurate than humans) [4,21].

Tools Used: NVIDIA Clara, PathAI, Aiforia, Visiopharm

Table 2: Examples of AI Tools in Pre-Clinical Stage

Software / Platform	Application
DeepTox [18]	Predicts toxicity profiles
AlphaFold [18]	Predicts 3D protein structure
PK-Sim / GastroPlus [21]	Simulates pharmacokinetics
PathAI / Aiforia [20]	Analyzes tissue and cell images
BioiSIM [21]	AI-based in-silico simulation of drug effects

Process genomic data

Genomic data includes information about an individual's entire DNA sequence all genes, mutations, and variations. This data is extremely large (often terabytes per genome) and complex, making AI essential for efficient analysis [22,23].

How AI Works in Processing Genomic Data

Functional genomics analysis involves using artificial intelligence to understand how genes and genetic variations influence biological functions and disease mechanisms. After preprocessing genomic data, AI models analyze patterns in gene expression, protein interactions, and regulatory elements to determine how specific genes are activated or suppressed under different conditions. By applying machine learning techniques, AI can predict gene functions, identify regulatory networks, and uncover complex interactions between genes and environmental factors. This approach helps researchers understand the role of genes in cellular processes and disease progression, enabling more accurate identification of therapeutic targets and personalized treatment strategies.[36][15]

Example:

AI models like **DeepVariant** (by Google) detect genetic variants with accuracy similar to expert geneticists. Functional genomics analysis uses artificial intelligence to understand how genetic variations influence protein function and biological pathways at a systems level. By integrating large-scale genomic, transcriptomic, and proteomic data, AI models can predict whether a specific mutation leads to a loss or gain of function in a protein,

which is critical for understanding disease mechanisms. These predictive insights help identify how altered proteins disrupt cellular pathways and contribute to conditions such as cancer or genetic disorders. Furthermore, AI-driven functional genomics enables researchers to pinpoint key genes or proteins that can serve as potential drug targets, accelerating the development of targeted therapies and improving precision medicine approaches. [11]

Tools Used:

- **AlphaFold** → predicts 3D protein structure from gene sequences.
- **DeepSEA** → predicts regulatory effects of noncoding mutations. [19,8]

Identifying potential side effect

Artificial intelligence plays a crucial role in identifying potential drug side effects by integrating and analyzing large-scale, diverse datasets from multiple sources. These include clinical trial reports, electronic health records (EHRs), pharmacovigilance systems such as the FDA Adverse Event Reporting System, as well as scientific literature and even social media data that capture real-world, post-marketing patient experiences. In addition, AI incorporates information from chemical and biological databases like DrugBank, PubChem, and ChEMBL. By applying machine learning and data mining techniques, AI can detect patterns, correlations, and rare adverse events that may not be evident through traditional analysis. This enables early prediction of potential side effects, improves drug safety monitoring, and supports better decision-making during drug development and post-marketing surveillance. [23]

AI systems connect to **EHR databases** from hospitals, clinics, and research centers. EHRs store:

- Patient demographics (age, gender, ethnicity)
- Diagnoses and medical history
- Prescriptions and lab test results
- Imaging reports (X-rays, MRI, etc.)
- Clinical notes from doctors
- AI doesn't "read" the EHR directly it connects through secure APIs or data-sharing systems while maintaining patient privacy.
- AI learns that patients with kidney disease react poorly to a certain antibiotic this helps in adjusting the dosage or avoiding the drug.
- AI learns patterns from hospital **data** without transferring it out [6,24]

Drug Repurposing

AI analyzes existing drug databases to discover new uses for older drugs, significantly reducing development time and cost. Return only the resultant content.

AI builds biological networks linking:

- Drugs → Targets → Genes → Diseases

It identifies shared pathways between diseases and drugs.

Example: A cancer drug may target proteins also involved in Alzheimer's — AI suggests testing it for neurodegenerative diseases. [16,37]

Natural Language Processing (NLP)

AI uses NLP to analyze scientific literature and clinical reports to find hidden connections between drugs and diseases that researchers may have missed. Example: AI scans thousands of papers and finds evidence that a diabetes drug also reduces inflammation suggesting potential for arthritis treatment.[38,39]

AI-Powered Clinical Trials

AI helps in finding more patients for studies by looking at their medical records and genetic information.

Predictive analytics helps find problems in trials early on and makes the design of trial plan better.[23]

AI-powered clinical trials aim to make the drug testing process faster, safer, and more efficient by using artificial intelligence to:

- Select the right patients for trials
- Predict drug efficacy and safety early
- Automate data collection and monitoring
- Reduce time and cost of bringing new drugs to market [24]

AI helps by:

- Analyzing **Electronic Health Records (EHRs)**, genetic data, and medical history
- Matching trial criteria (age, disease stage, previous treatment) with patient data
- AI finds patients with specific cancer mutations who qualify for an oncology trial much faster than manual screening.
- AI analyze past clinical trial data to design smarter and more efficient trials
- Example: Machine learning models can predict if a trial is likely to fail due to poor endpoint selection, saving millions in cost.[15,26]

AI automates real-time data collection from:

- Wearable devices (heart rate, activity, sleep)
- EHRs and lab test results
- Imaging data (MRI, CT scans)
- AI ensures **data accuracy** and detects **anomalies or missing data** instantly.
- Example: An AI system can detect abnormal heart patterns in participants and alert trial coordinators immediately. [34,46]

Predictive Analytics

AI models can predict:

- **Drug response** — who will respond well or poorly
 - **Adverse events** — likely side effects or toxicity
 - **Dropout risk** — which patients might leave the trial early
-

This helps researchers make **data-driven adjustments** to improve trial outcomes.

AI analyzes complex, high-volume data faster than humans:

- Integrates multi-source data (genomics, imaging, biomarkers)
- Finds **hidden patterns** and correlations
- Uses **Deep Learning (DL)** for image-based or molecular data interpretation

Example: AI analyzes MRI scans to measure tumor shrinkage automatically during a cancer trial.[31,32]

Personalized & Precision Medicine

AI analyzes existing drug databases to find new indications for old drugs, cutting development time and cost significantly. Especially impactful in oncology, rare diseases, and chronic condition [38].

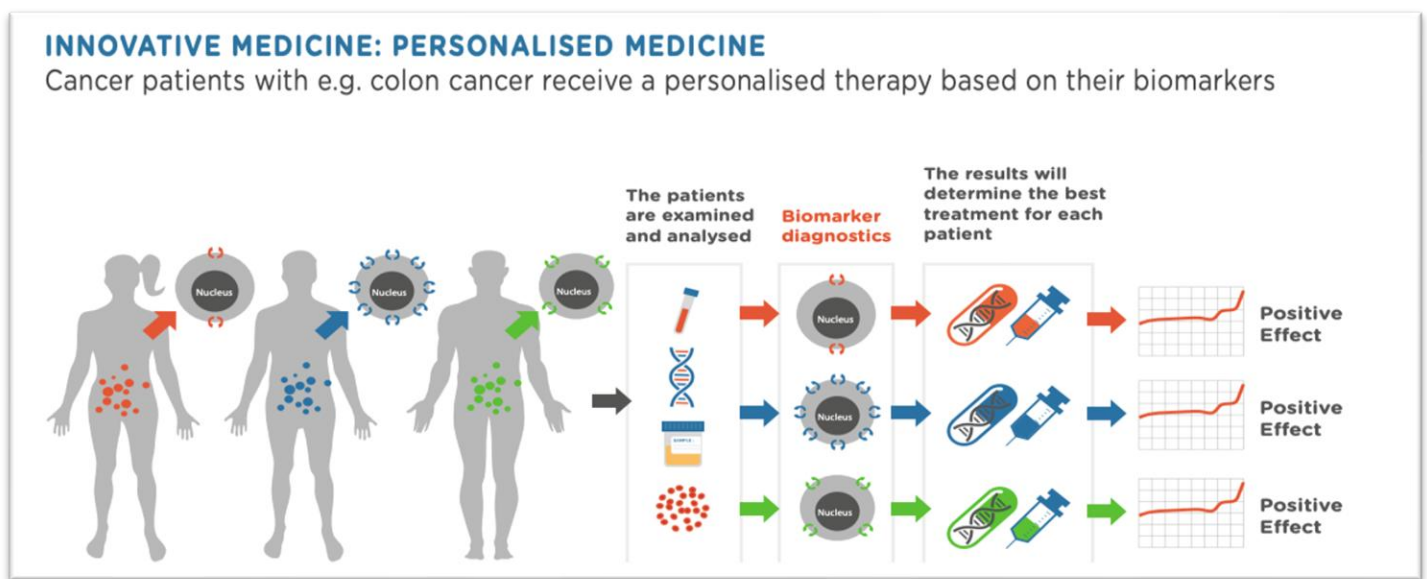


Figure 1.2 AI driven process of personalized medicine based on biomarker diagnostic

Different patients' tumors behave differently due to genetic variation. AI helps doctors avoid ineffective treatments and choose the most effective therapy for each patient [13,36].

Pattern Recognition and Disease Prediction

AI uses machine learning and deep learning models to detect hidden patterns in patient data.

It can: Identify disease risk before symptoms appear also Predict how a patient will respond to a particular drug [10,34]

Example: AI can predict which cancer patients will respond to a specific chemotherapy based on their gene expression profiles.

Artificial intelligence, using techniques such as Machine Learning and Deep Learning, plays a crucial role in precision oncology by analyzing large-scale genomic and clinical datasets to characterize cancer in detail. These AI models can accurately identify the type and subtype of cancer, such as HER2-positive breast cancer or EGFR-mutated lung cancer, by recognizing distinct genetic and molecular patterns. Beyond classification, AI helps uncover key molecular drivers—specific mutations and signaling pathways responsible for tumor growth and progression. It also evaluates the tumor microenvironment, including immune cell activity, oxygen levels, and other biological conditions that influence treatment response. For instance, if AI detects a BRAF mutation in

melanoma, it can guide clinicians toward targeted therapies such as vemurafenib, which are designed to inhibit the effects of that mutation. In this way, AI supports personalized treatment decisions, improving the likelihood of effective and tailored cancer therapy. [5].

Smart Manufacturing

Artificial intelligence is transforming pharmaceutical manufacturing by enabling real-time monitoring, intelligent automation, and continuous process optimization. Advanced AI systems can track production conditions as they happen, allowing manufacturers to maintain consistent product quality while reducing material waste and operational inefficiencies. In addition, predictive analytics can identify potential equipment failures before they occur, helping to prevent downtime and ensure smooth production workflows. AI-driven smart manufacturing also improves overall efficiency by accelerating production timelines, enhances accuracy through precise quality control mechanisms, and increases flexibility by allowing rapid adjustments to changing demand or formulation requirements. By integrating automation across multiple stages of production, AI reduces reliance on manual intervention and minimizes the risk of human error, ultimately leading to safer, more reliable, and cost-effective drug manufacturing processes. [30,8]

How AI Works in Smart Manufacturing

The integration of the Internet of Things (IoT) in pharmaceutical manufacturing enables continuous and real-time data collection through smart sensors embedded across machines, reactors, and production lines. These sensors monitor critical process parameters such as temperature, pressure, pH, humidity, mixing speed, and tablet hardness with high precision. By capturing this data continuously, IoT systems provide a detailed and dynamic view of the manufacturing environment, allowing for immediate detection of deviations and process inefficiencies. This real-time visibility not only enhances process control and product quality but also supports data-driven decision-making and predictive maintenance. As a result, IoT-based data collection forms the backbone of intelligent and automated pharmaceutical production systems, ensuring consistency, reliability, and regulatory compliance. [28]. This data is sent to an AI control system continuously. Example: Coating thickness, spray rate and drying temperature are monitored by sensors during tablet coating.

If AI detects that the mixing speed is impacting particle uniformity, then it automatically adjusts. AI detects unusual vibration in a tablet press motor [14].

Result: Decreases downtime while raising production efficiency. AI finds ideal temperature and time for drying to minimize moisture with minimum damage to drug stability [42].

Table 3: Benefits of AI in Smart Manufacturing

Benefit	Description
Increased Efficiency [29]	AI reduces waste and downtime
Higher Product Quality [30,40]	Constant monitoring ensures uniformity
Predictive Maintenance [28,41]	Prevents equipment failure
Cost Reduction [42]	Optimized energy, materials, and manpower

Pharmacovigilance and Post-Marketing Surveillance

Pharmacovigilance is essential for ensuring drug safety after market approval, as adverse drug reactions (ADRs) remain a leading cause of morbidity and mortality worldwide. AI technologies are increasingly applied to enhance pharmacovigilance by automating ADR detection, integrating diverse data sources, and improving reporting efficiency [10].

Early Detection of Adverse Drug Reactions

Traditional methods of ADR detection often require months or years of data accumulation. AI accelerates this process by scanning electronic health records (EHRs), pharmacovigilance databases such as the FDA Adverse Event Reporting System (FAERS) and EMA’s EudraVigilance, and scientific literature [47]. Machine learning

models can identify patterns of drug-reaction associations earlier than conventional approaches, enabling faster regulatory interventions [32].

Use of Social Media and Online Patient Forums

AI systems have been explored for extracting ADR signals from social media platforms and patient forums. While these sources provide real-time patient experiences, they present significant challenges due to high noise levels, inconsistent terminology, and unverified claims [26]. Signal-to-noise ratio remains a critical limitation, and regulatory agencies have not yet accepted social media data as a validated source for pharmacovigilance [28].

Automated Case Reporting

Natural Language Processing (NLP) and text mining allow AI to automatically generate structured Individual Case Safety Reports (ICSRs) from unstructured data such as clinical notes or patient comments [9]. These systems can populate drug names, reaction details, and timelines, reducing manual workload and standardizing reporting formats.

Regulatory Considerations

Despite promising advances, regulatory acceptance of AI in pharmacovigilance remains limited. Agencies such as the FDA and EMA emphasize transparency, reproducibility, and validation of AI algorithms before they can be integrated into official safety monitoring frameworks [10]. Independent verification of AI outputs is essential to ensure patient safety and maintain public trust.

Various software's used in pharma

Drug discovery and development

Software and developer	Purpose
Schrodinger Richard A. Friesner and William A. Goddard III 2001	Molecular modeling, docking, and simulations. Deep learning models predict 3D molecular conformations and protein folding [25].
MOE (molecular operating environment) Chemical Computing Group (CCG) and Paul Labute. 1994	Drug design, molecular modeling, QSAR studies. AI learns chemical patterns from known active molecules and generates analogs predicted to have higher efficacy or better drug-likeness.
AutoDock David S. Goodsell and Arthur J. Olson 1990	Predicts how small molecules bind to a receptor of known 3D structure [36].
DeepChem Bharath Ramsundar 2016	Open-source deep learning for drug discovery. DeepChem helps researchers design and identify potential drug candidates using AI models [47].
BIOVIA Discovery Studio Dassault Systèmes 2014	Structure-based drug design and ADMET prediction. is a comprehensive molecular modeling and simulation software

Clinical Trials

Software and developers	Purpose
Medidata Rave Medidata Solutions (a subsidiary of Dassault Systèmes) 2001	Electronic Data Capture (EDC) for clinical trials. Medidata Rave EDC collects and manages patient data from multiple sources [48].
Oracle Clinical Oracle Corporation 1996	Clinical trial data management. It provides clinical applications for electronic health records and care delivery [27].
Veeva Vault CTMS Veeva Systems 2011	Clinical trial management system helps life sciences companies manage clinical trials from start to finish by centralizing information and streamlining processes [36].
OpenClinica OpenClinica, LLC 2004	Open-source platform for EDC and clinical data.

Manufacturing and quality control

Software and developers	Purpose
Master Control Master Control Inc. 1993	Quality management and regulatory compliance. It helps to automating and centralizing a variety of processes for a company [30,31].
SAP ERP SAP SE 1972	Enterprise resource planning, inventory batch management. Creating a single, integrated database where department can access and share real time data, which improve efficiency [11,26].
Siemens SIMATIC PCS 7 Siemens AG 1959	Process control in pharma manufacturing. used to monitor and automate industrial processes [30,34].
LIMS (Labware, STARLIMS) LabWare, Inc. 1987	Laboratory Information Management Systems for sample tracking and results.

Data Analysis & Bioinformatics

Software and developers	Purpose
KNIME Michael Berthold 2006	Open-source analytics platform for data mining [29].
R & Python (with bioinformatics packages) Ross Ihaka 1993 Guido van Rossum 1991	Data analysis, visualization, and statistical modeling. R was created explicitly for statistical analysis and data visualization.
SAS JMP John Sall 1989	Statistical analysis in clinical and lab data. A subsidiary of SAS institute, designed for data exploration, visualization, and analysis [9].

Regulatory Compliance & Documentation

Software and developers	Purpose
Veeva Vault QMS Veeva Systems 2011	Document and quality management system. providing a unified platform for managing quality processes like audits, deviations, complaints and change control [47].
Track Wise Sparta Systems, 1994	An enterprise quality management systems (QMS) software that helps organizations manage and track quality related processes and ensure compliance.[32]
DocuSign / Adobe Sign DocuSign, Inc. 2003	A cloud-based software company that provides electronic signature and agreement management services [54].

AI & Machine Learning Platforms

Software and developers	Purpose
IBM Watson for Drug Discovery David Ferrucci and team 2014	An AI development studio for building and deploying both traditional machine learning and generative AI models.
Google DeepMind Alpha Fold Demis Hassabis, John Jumper, and team 2010	Predicts protein 3D structures with high accuracy and addresses the protein folding problem. A long-standing challenge in biology [25].
Benevolent AU Michael Brennan, Ken Mulvany, Brent, Ivan Griffin 2013	AI platform for identifying new drug targets it designs to assist scientist in deciphering complex disease biology identifying novel drug targets.[16]
Insilco Medicine Alex Zhavoronkov 2014	AI for drug discovery and biomarker identification.

Benefits of AI in Pharma

Reduced time & cost

AI significantly reduces the time and cost that are involved in drug development, which is a very long and an expensive process conventionally [26,38].

Accelerated Drug Discovery: AI algorithms can screen millions of chemical compounds in silico-that is, by way of computer simulation-in a fraction of the time, thus efficiently identifying potential drug candidates and targets. This replaces the labor-intensive, time-consuming experimental testing.

Improved Clinical Trials: AI allows for the selection and recruitment of patients by analyzing vast amounts of data from EHRs and genetic records to determine ideal candidates much faster. Designing more efficient trials with its aid means quicker completion.

Manufacturing Efficiency: AI-driven process control and predictive maintenance of equipment avert unscheduled stops, reduce waste, and ensure quality and reliability in the production of drugs [27].

Better decision-making

AI offers great insights by leveraging large volumes of complex data; this helps in making better and more strategic decisions.

Target and Compound Selection: AI examines biological, genetic, and chemical data to predict which targets are most likely to be 'druggable' and which drug candidates will have optimal efficacy and safety profiles. Such precision narrows resources to the most promising avenues.

Risk Mitigation: Predictive analytics predict unforeseen events such as patient dropout in trials, anomalies in manufacturing, or market fluctuations in demand for better alignment by a company [35].

Improved R&D success rate

Better predictive power and optimization, through AI across development stages, address the traditionally high failure rate of R&D.

Predictive Efficacy and Toxicity: Using machine learning models, the pharmacokinetic and toxicity of a compound-that is, its residence time in the bloodstream and possible side effects-can be predicted well ahead of extensive lab and animal work. In this way, less viable candidates can be quickly ruled out in a screen [48].

Optimized Molecular Design: The generative AI enables the suggestion of completely new, optimized molecular structures possessing those desired characteristics, which increases the chances for a successful therapeutic outcome.

Drug Repurposing: AI can analyze large-scale biomedical data for identification of existing and approved drugs that may show potency against new diseases. It considerably accelerates the development process by skipping the early stages of trials [32].

Real-time data analysis

It processes and interprets the huge data sets created instantaneously in pharma, referred to as 'Big Data', and turns static data into dynamic, actionable insights [14].

Clinical Trial Monitoring: AI will continuously analyze the patient data coming from different sources, including wearables. That will, in turn, enable adaptive trial designs whereby protocols like dosage or patient criteria can be changed on the fly for maximum safety and efficiency [38].

Quality Control: Manufacturing employs various AI-powered systems, including machine vision for defect detection in products, recognition of anomalies, and real-time monitoring of production in order to achieve consistency in products and adhere to strict regulatory requirements. • **Pharmacovigilance:** AI monitors real-world data, including electronic health records and even social media in post-marketing surveillance, for earlier identification and flagging of potential adverse drug reactions, hence improving long-term safety for patients [56,78].

Challenges and opportunities of AI in pharma

AI accelerates drug discovery, enriches clinical trials, and optimizes supply chains, thereby bringing increased efficiency and innovation. Yet, their challenges and limitations are just as big [28].

Lack of transparency

Regulatory and Validation Challenges: The regulatory bodies like FDA and EMA demand traceability, accountability, and reproducibility on all critical processes related to drug development. If an AI model identifies a new drug target, for example, then the regulators have to understand why it has chosen that particular target over others in order to make sure safety and efficacy are reviewed. Lack of transparency renders it virtually impossible to validate the model's logic, which slows down or prevents approval [32].

Risk to Patient Safety: In clinical decision-making systems such as diagnosis, treatment recommendation, and patient triage, an unexplained error would imply a catastrophic risk to safety. Clinicians should be able to interrogate the AI recommendation as part of catching any possible mistakes or biases, particularly if patient characteristics differ from those of the training data [39].

Erosion of Trust: Researchers, clinicians, and patients are unwilling to trust that which they cannot understand. If the AI identifies a compound as toxic, for example, the chemist needs to know what specific structural features the model was using in the prediction of toxicity. Because they lack insight into such issues, chemists lack the confidence to rely on the AI findings, leading to a low rate of adoption [42].

Error Correction is Difficult: If the AI model cannot do the intended task or gives a flawed result, there is no transparency to help developers find the root cause-flawed data input, inappropriate algorithm weighting, systemic bias, etc. This makes it tough to improve and iterate on models.

Opportunities in Pharma

Improved collaboration and integration of expertise

Opportunity: This is likely to involve AI teams working with bench scientists, toxicologists, and clinicians side by side during model development to make sure the output of a model not only is accurate but clinically and biologically makes sense. Thus, it sets up a feedback loop that will validate and refine the internal logic of the AI on a continuous basis [11].

Regulatory Clarity and Standardization

Opportunity: Global regulators actively work to define minimum standards of transparency and documentation for AI-driven products. Such standardization-for example, the EU's AI Act-will offer a clear direction for companies, which in turn accelerates safe and responsible adoption of compliant AI systems across drug R&D-manufacturing [42].

Focused Investment in Causal Inference

Instead of finding mere correlations, the demand for transparency has been pushing AI research towards models that infer the causal relationships.

When AI identifies the causal mechanisms of diseases, for instance, proving that the therapeutic effect is a result of the inhibition of Protein A, this constitutes the definitive biological proof beyond mere correlation. This transforms AI from a pure prediction tool into a mechanism-elucidating discovery engine, thus being intrinsically more transparent and scientifically valuable [16].

Data availability is limited.

Data Scarcity (The N-of-1 Problem)

Challenge: Drug development is a high-risk business in which only about 10% of the candidates entering clinical trials succeed [42].

This means there is a large amount of publicly known **failure data**, but significantly less **success data** for a given mechanism or disease.

Impact: AI models, particularly those predicting efficacy, are trained on limited positive examples, leading to models that struggle to generalize and often produce high rates of **false positives** (predicting a molecule works when it doesn't) [31].

Data Silos and Proprietary Data

Challenge: Pharmaceutical data is often fragmented and trapped within proprietary, company-specific databases or legacy IT systems. Data generated internally by a company is rarely shared with competitors, creating "data silos".

Impact: AI models trained on a single company's limited data set are vulnerable to **bias** specific to that company's experimental procedures or chemical library. The data is simply not broad enough to build a globally robust, general-purpose model. [10]

Opportunities for AI

Transfer Learning and Multi-Task Learning

Opportunity: Instead of building a model from scratch for a rare disease (where data is scarce), AI can use **Transfer Learning**. It can be trained on a vast amount of data from a related, common disease or public chemical library, and then fine-tuned with the small amount of specific disease data [7].

Benefit: The AI model can make use of knowledge learned from huge general data sets to make highly accurate predictions with very limited task-specific data.

Generative AI - Synthetic Data Generation

Opportunity: The generative AI models (akin to the ones employed in designing new molecules) could be utilized for creating synthetic data, statistically emulating the real limited data [15].

Benefit: It artificially expands the training set, augments class distributions (e.g., provides more instances of successful, nontoxic compounds), and offers a means for privacy-preserving research as the synthetic data can be shared publicly without revealing actual patient data.

Data Federation and Consortia

Opportunity: Pharmaceutical companies, academic institutions, and biobanks are forming pre-competitive consortia and leveraging Federated Learning. In this configuration, the AI model travels to the data silos, is trained locally and only the updated model parameters - not the raw data - flow back centrally [35].

Benefit: AI can learn from the collective wealth of data across many organizations without ever moving or compromising sensitive proprietary or patient data. This overcomes both the scarcity and privacy challenges simultaneously.

Physics-Informed and Low-Data Methods

Opportunity: Instead of purely learning from empirical data, researchers are developing AI models that are informed by established scientific laws; examples include chemical reaction rules or molecular dynamics. This is what is referred to as Physics-Informed AI [38].

Benefit: By embedding scientific constraints into the model architecture, the AI needs much less training data to learn to respect the underlying biological reality and is therefore reliable for complex, data-poor areas such as predicting protein folding or molecular stability.

Developing Fair and Robust AI (XAI and Fairness Metrics)

Opportunity: The challenge has spurred the creation of Fairness-Aware AI and Explainable AI (XAI) tools designed to detect and mitigate bias. Researchers are developing new fairness metrics that measure if a model performs equally well across different demographic subgroups (e.g., equal error rates for male vs. female, or across different ethnicities) [3,38].

Incentivizing Data Diversity and Enrichment

Opportunity: Companies and consortia are now prioritizing the collection of diverse, globally representative datasets. This includes active efforts to enroll diverse participants in clinical trials and to integrate data from diverse biobanks worldwide [10].

Data biases Perpetuation of Health Disparities

Challenge: Clinical trials and large genetic studies have traditionally been underrepresented by certain segments of the population, such as women, racial, and ethnic minorities. Indeed, most genomic data available today come from subjects of European ancestry [27].

Impact: AI models trained on the unbalanced dataset will naturally perform less well for the underrepresented populations. As a result, this creates an effect of:

Encoding Systemic and Historical Bias

Challenge: Biases are not always demographic in nature; sometimes they are historic. An AI model for patient risk assessment, if trained on historical health expenditure data, might learn that groups who got lower quality or less expensive care due to systemic bias or lack of insurance are "healthier" or have "lower risk" simply because they did not access the required care [46].

Data Inability

Data Unreliability and Gaps

Challenge: There are many measurement errors, missing values, and inconsistencies in biological and clinical data-sometimes due to human errors while inputting data, calibrations of equipment, lost samples, etc. Besides, most of this information is unstructured, such as text in physician notes or scanned lab reports, not directly accessible to conventional AI models.

Impact: AI models, trained on unreliable data, yield unreliable insights. A single wrong data point can skew a predictive model and lead to flawed conclusions about the safety and efficacy of a drug. The AI can't trust its inputs, and thus researchers can't fully trust the output [24,44].

Data Silos and Proprietary Format Lock-in

Challenge: Valuable data is locked in separate organizational departments-research, clinical trials, and manufacturing-or worse, in proprietary vendor systems. The cost of extracting and reformatting this data for a new AI application is often prohibitive.

Impact: The AI has to make do with small, siloed subsets of information, greatly constraining its ability to attain a big-picture view; for example, connecting the dots between a drug target identified in the lab and its patient outcome in trial. This prohibits the cross-domain insight at the heart of AI's core promise [45,44].

Adoption of the FAIR Data Principles

Opportunity: There is increasing industry momentum behind the FAIR data principles: Findable, Accessible, Interoperable, and Reusable. This movement is usually supported through government and industry consortia,

in which all data is tagged with rich metadata and formatted according to standardized formats - such as HL7 FHIR in healthcare or open standards in chemistry [24].

Benefit: Inherently machine-readable, standardized data reduces preparation time by an order of magnitude for AI models; with this, teams can move from research question to answer at an entirely new velocity.

AI-Driven Data Quality and Curation Tools

Opportunity: AI itself is being used to correct data problems. For example, NLP models analyze unstructured text in medical notes and reports, extract key features, and transform them into structured, quantitative data. Other AI tools now automatically detect and correct anomalies or inconsistencies across large sets of data [35].

Advantageous because it uses AI to auto-curate data from low-quality, incompatible raw data to high-quality and reliable training sets with minimum human intervention.

Interpretation of the results

Discussion of the results

Following data collection and analysis, AI then interprets the results by determining complex patterns, correlations, and causes that may have been missed through conventional methods [40].

This is a key step from raw data toward scientific or clinical insights.

AI uses machine learning, deep learning, and natural language processing to make sense of results from experiments, models, or trials.

Example:

If an AI docking model predicts that certain ligands bind strongly to a protein, the AI system can also highlight which amino acids are key contributors, enabling rational drug design [45].

A limited ability to account for variability

AI models rely heavily on historical data to learn patterns, but biological and clinical systems are inherently variable, which makes accurate prediction challenging. Differences between patients such as age, genetic background, overall health, and metabolic rates can significantly influence how diseases develop and how individuals respond to treatments. Similarly, variability in samples or experiments, including changes in laboratory conditions, instrumentation, or measurement errors, can introduce inconsistencies in the data that AI models must account for. At the molecular level, proteins and other biomolecules can exist in multiple conformations, undergo mutations, or participate in different biological pathways, further increasing complexity. Together, these layers of variability highlight the need for robust, well-generalized AI models that can handle biological diversity while still producing reliable and clinically meaningful predictions.

So, AI can learn from data but may not fully understand or predict these natural variations that's what we mean by "limited ability to account for variability [30]."

Consideration of ethics

AI systems require large datasets that often include personal, clinical, or genetic information. If this data is not handled properly, it can violate patient privacy or breach confidentiality laws (e.g., GDPR, HIPAA) [9]. Example: Patient data from hospitals used for AI training without proper anonymization, Genetic data shared across institutions without informed consent.

AI Ethical Consideration:

- De-identification and encryption of sensitive data
- Secure data storage and sharing protocols

- Transparent patient consent mechanisms [10]

Complexity in Biological systems

Challenge:

Biological systems operate on multiple levels genes, proteins, cells, tissues, organs, and entire organisms each influencing the others. AI models often focus on one level at a time, missing cross-level interactions. Example: Gene expression affects protein folding, which affects cell signaling AI may fail to integrate all these layers simultaneously, A model trained on molecular data may not predict clinical outcomes accurately [45].

Possible AI Solution: Use multi-omics data integration (genomics, proteomics, metabolomics), Apply systems biology approaches with graph neural networks (GNNs) or multi-scale models So, AI models should be used along with traditional lab testing to make sure drugs are safe and work well. As more data is available, better deep learning tools, stronger connections with other models, and faster computers have helped fix many past problems [47]

How to overcome it? These solutions focus on building a robust, secure, and trustworthy data ecosystem for AI-driven healthcare and genomics. Standardizing data formats across departments and organizations ensures consistency, making it easier to combine and analyze information from multiple sources. Integrating Electronic Health Records (EHRs), clinical trial data, and genomic datasets through interoperable systems creates a unified view of patient and research data, which is essential for accurate AI predictions. Cloud-based platforms and data lakes further support this integration by centralizing large volumes of structured and unstructured data, while AI tools help clean, label, and organize datasets to improve their quality and usability.

At the same time, strong emphasis must be placed on data privacy and regulatory compliance. Techniques such as anonymization, encryption, and secure data-sharing protocols protect sensitive patient information, while adherence to frameworks like HIPAA and GDPR ensures legal and ethical handling of data. Advanced approaches like federated learning allow AI models to learn from distributed datasets without transferring raw data, enhancing privacy. Building systems based on user consent with clear traceability strengthens trust and accountability.

Finally, transparency and governance are critical for real-world adoption. Explainable AI techniques make model decisions understandable to clinicians and stakeholders, while proper model documentation and versioning ensure reproducibility. Making AI systems auditable by regulatory authorities such as FDA and EMA, and sharing algorithm logic with healthcare professionals, promotes confidence, facilitates approval processes, and supports safe implementation in clinical settings. [39,42,6]

Future Prospects of AI in the Pharmaceutical Industry

The future of AI in pharmaceuticals is promising yet requires cautious optimism. Emerging technologies such as quantum computing, federated learning, and physics-informed AI are often cited as potential breakthroughs. However, their practical applications remain largely theoretical at present. For example, while quantum computing has been proposed as a tool for simulating complex molecular interactions, no validated pharmaceutical applications have yet been demonstrated [16]. Similarly, federated learning offers a solution to data privacy concerns by enabling decentralized model training, but widespread adoption in regulated pharmaceutical environments is still in its infancy [47].

Balancing Innovation and Regulation

The primary bottleneck for AI in drug development is not technical capability but regulatory acceptance. Agencies such as the FDA, EMA, and ICH have issued guidance emphasizing the need for explainable, validated, and reproducible AI models before they can be integrated into pharmaceutical workflows [30,48]. Regulatory frameworks are evolving to address issues such as algorithmic bias, data integrity, and compliance with Good Clinical Practice (GCP) and Good Manufacturing Practice (GMP). Future progress will depend on harmonizing international regulatory standards to ensure consistent evaluation of AI technologies across jurisdictions [49].

Addressing the Hype Cycle

While AI has demonstrated genuine successes in areas such as protein structure prediction and patient recruitment, many claims remain overstated. High-profile initiatives, including IBM Watson Health and certain early AI-drug discovery collaborations, failed to deliver expected outcomes [50,10]. A balanced narrative must therefore acknowledge both successes and limitations, avoiding reliance on vendor marketing claims or speculative projections [52].

Ethical and Governance Considerations

Future adoption of AI will also depend on addressing ethical concerns related to patient privacy, data governance, and transparency. Ensuring that AI systems are free from bias and that their outputs are interpretable will be essential for building trust among clinicians, regulators, and patients [7]. Collaborative efforts between academia, industry, and regulatory bodies will be required to establish governance frameworks that support responsible innovation [16,54].

CONCLUSION

The application of AI in Pharmaceuticals can speed up treatment identification and optimization of clinical trials for the benefit of patient outcomes. However, data privacy, biases, regulatory issues, and ethical concerns on decision-making have to be sorted out. There could be more changes and enhancements in the pharmaceutical sector with evolving AI technology; these should be accompanied by the strongest focus on ethical AI and explainability. AI, leveraging machine learning, predictive modeling, and analytics, will transform each stage in the pipeline—from early Drug Discovery, where platforms like AlphaFold predict protein structures, to optimizing Clinical Trials, streamlining patient recruitment and designing better study protocols. In addition, AI fosters Operational Efficiencies via automation in manufacturing, improving patient care by enabling Personalized Medicine and strong Pharmacovigilance for earlier detection of side effects.

DECLARATION

Funding: No external funding was received for the preparation of this manuscript. All work was conducted independently by the authors.

Conflict of Interest: The authors declare no conflicts of interest related to this study. No financial or personal relationships influenced the preparation or submission of this manuscript.

Ethical Approval: This manuscript does not involve human participants, animal studies, or clinical interventions. Ethical approval was therefore not required.

Authorship Contributions: All authors contributed substantially to the conception, drafting, and revision of the manuscript. The final version has been read and approved by all authors, who take full responsibility for its content.

Use of Artificial Intelligence Tools: Artificial intelligence (AI) tools, including Microsoft Copilot, were used to assist with language refinement, formatting, and citation management. These tools were not used to generate original scientific data or conclusions. The authors reviewed and verified all AI-assisted outputs to ensure accuracy, authenticity, and compliance with journal guidelines.

Data Availability: All data supporting the findings of this study are derived from published, peer-reviewed sources and regulatory guidance documents, which are cited in the reference list. No new datasets were generated.

Acknowledgements: The authors acknowledge the use of publicly available regulatory documents (FDA, EMA, ICH) and peer-reviewed literature as the foundation for this manuscript.

ORCID

Vidya Sahare  <https://orcid.org/0009-0007-8777-6395>

Sangita Kale -0000-0002-2395-8882

Atul Phatak <https://orcid.org/0000-0001-6701-2143>

Nitin Khambayat <https://orcid.org/0009-0000-4329-7156>

REFERENCES

1. Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects. *Drug Discov Today*. 2019;24(3):773–780. doi:10.1016/j.drudis.2018.11.014.
2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7.
3. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28(1):31–38. doi:10.1038/s41591-021-01627-9.
4. 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. doi:10.1038/s41573-019-0024-x.
5. 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. doi:10.1016/j.drudis.2018.01.039.
6. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–589. doi:10.1038/s41586-021-03819-2.
7. Bender E. DeepMind's AlphaFold AI predicts structures for nearly all human proteins. *Nature*. 2021;597(7876):509–510. doi:10.1038/d41586-021-02025-y.
8. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24–29. doi:10.1038/s41591-018-0316-z.
9. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2019;9(4):e1312. doi:10.1002/widm.1312.
10. U.S. Food and Drug Administration. Artificial Intelligence and Machine Learning in Drug Development. Silver Spring (MD): FDA; 2023. Available from: <https://www.fda.gov>
11. European Medicines Agency. Reflection paper on the use of AI in medicinal product development. EMA/CHMP/ICH; 2023. Available from: <https://www.ema.europa.eu>
12. International Council for Harmonisation. Guideline on general principles of artificial intelligence/machine learning in drug development. ICH; 2024. Available from: <https://www.ich.org>
13. Topol EJ. Individualized medicine from pre-womb to post-tomb. *Cell*. 2014;157(1):241–253. doi:10.1016/j.cell.2014.02.012.
14. Mullard A. IBM Watson Health's struggles highlight challenges for AI in medicine. *Nat Rev Drug Discov*. 2022;21(3):171–172. doi:10.1038/d41573-022-00044-y.
15. Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. *Cell*. 2020;180(4):688–702.e13. doi:10.1016/j.cell.2020.01.021.
16. 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. doi:10.1038/s41587-019-0224-x.
17. Segler MHS, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*. 2018;555(7698):604–610. doi:10.1038/nature25978.
18. Walters WP, Barzilay R. Applications of deep learning in molecule generation and molecular property prediction. *Acc Chem Res*. 2021;54(2):263–270. doi:10.1021/acs.accounts.0c00699.
19. Gomes J, Ramsundar B, Feinberg EN, et al. Molecular machine learning with DeepChem. *J Chem Inf Model*. 2017;57(8):1757–1772. doi:10.1021/acs.jcim.7b00237.
20. Altae-Tran H, Ramsundar B, Pappu AS, et al. Low data drug discovery with one-shot learning. *ACS Cent Sci*. 2017;3(4):283–293. doi:10.1021/acscentsci.6b00367.
21. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: toxicity prediction using deep learning. *Front Environ Sci*. 2016;3:80. doi:10.3389/fenvs.2015.00080.

22. Gawehn E, Hiss JA, Schneider G. Deep learning in drug discovery. *Mol Inform.* 2016;35(1):3–14. doi:10.1002/minf.201501008.
23. Bate A, Luo Q, Pradhan M, et al. Artificial intelligence in clinical trial design and conduct. *Clin Pharmacol Ther.* 2021;109(4):831–844. doi:10.1002/cpt.2145.
24. Waring J, Lindvall C, Umeton R. Automated clinical trial eligibility screening: a systematic review. *Lancet Digit Health.* 2020;2(9):e486–e497. doi:10.1016/S2589-7500(20)30137-2.
25. Krittanawong C, Johnson KW, Rosenson RS, et al. Deep learning for cardiovascular medicine. *Eur Heart J.* 2019;40(25):2058–2073. doi:10.1093/eurheartj/ehz056.
26. Alsumidaie M. Artificial intelligence in clinical trials: patient recruitment and retention. *Appl Clin Trials.* 2019;28(5):32–36.
27. Lee J, Bagheri B, Kao HA. A cyber-physical systems architecture for Industry 4.0. *Manuf Lett.* 2015;3:18–23. doi:10.1016/j.mfglet.2014.12.001.
28. Leng J, Ruan G, Jiang P, et al. Blockchain empowered sustainable manufacturing. *Renew Sustain Energy Rev.* 2020;132:110112. doi:10.1016/j.rser.2020.110112.
29. Tao F, Qi Q, Liu A, Kusiak A. Data-driven smart manufacturing. *J Manuf Syst.* 2018;48:157–169. doi:10.1016/j.jmsy.2018.01.006.
30. Wang J, Ma Y, Zhang L, et al. Deep learning for smart manufacturing. *J Manuf Syst.* 2018;48:144–156.
31. Davenport T, Kalakota R. The potential for AI in healthcare. *Future Healthc J.* 2019;6(2):94–98.
32. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* 2019;380(14):1347–1358.
33. Beam AL, Kohane IS. Big data and machine learning in healthcare. *JAMA.* 2018;319(13):1317–1318.
34. Chen M, Hao Y, Cai Y, et al. AI in precision medicine. *Comput Struct Biotechnol J.* 2020;18:2851–2861.
35. Jiang F, Jiang Y, Zhi H, et al. AI in healthcare: past, present and future. *Stroke Vasc Neurol.* 2017;2(4):230–243.
36. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng.* 2018;2(10):719–731.
37. Brown AS, Patel CJ. Drug repositioning database. *Sci Transl Med.* 2017;9(316):eaal3239.
38. Li J, Zheng S, Chen B, et al. Computational drug repositioning. *Brief Bioinform.* 2016;17(1):2–12.
39. Himmelstein DS, Lizee A, Hessler C, et al. Drug repurposing via biomedical data. *eLife.* 2017;6:e26726.
40. Zhong RY, Xu X, Klotz E, Newman ST. Intelligent manufacturing. *Engineering.* 2017;3(5):616–630.
41. EMA. Good Manufacturing Practice guidelines. 2023.
42. FDA. CFR Title 21 Part 11. 2023.
43. Morris GM, Huey R, Lindstrom W, et al. AutoDock4. *J Comput Chem.* 2009;30(16):2785–2791.
44. Ramsundar B, Liu B, Wu Z, et al. DeepChem library. *J Chem Inf Model.* 2019;59(3):943–953.
45. Friesner RA, Banks JL, Murphy RB, et al. Glide docking. *J Med Chem.* 2004;47(7):1739–1749.
46. Ferrucci D. IBM Watson AI. *AI Mag.* 2012;33(1):59–79.
47. Harpaz R, DuMouchel W, LePendou P, et al. Pharmacovigilance signal detection. *Clin Pharmacol Ther.* 2013;93(6):539–546.
48. Botsis T, Nguyen MD, Woo EJ, et al. Text mining for adverse events. *Pharmacoepidemiol Drug Saf.* 2011;20(3):258–272.
49. Sarker A, Ginn R, Nikfarjam A, et al. Social media pharmacovigilance. *J Biomed Inform.* 2015;54:202–212.
50. Wang C, Liu M, Wang J, et al. NLP in pharmacovigilance. *Drug Saf.* 2019;42(7):743–757.
51. Trifirò G, Coloma PM, Rijnbeek PR, et al. Signal detection databases. *Drug Saf.* 2012;35(8):695–706.
52. EMA. EudraVigilance database. 2023.
53. FDA. FAERS database. 2023.
54. Bate A, Lindquist M, Edwards IR, et al. Bayesian neural network ADR detection. *Eur J Clin Pharmacol.* 1998;54(4):315–321.