Journal
Of the
Oriental Institute
M.S. University of Baroda

ISSN: 0030-5324
UGC CARE Group 1

IMPROVISING DRUG DEVELOPMENT PROCESS USING AI/ML TOOLS AND METHOD: A REVIEW PAPER

Pooja Mehta[1], Manish Shrimali[2]

[1,2]Department of Computer Science,
Janardan Rai Nagar Rajasthan Vidhyapeeth, Udaipur, Rajasthan, India.

ABSTRACT

We cannot simply accept that testing new drugs will continue to be a slow and expensive process. AI has the potential to disrupt the current approach to clinical trials — from patient recruitment to adherence monitoring and data collection – and it is time to seize these opportunities.
future of clinical development is on the verge of a major transformation due to convergence of large new digital data sources, computing power to identify clinically meaningful patterns in the data using efficient artificial intelligence and machine-learning algorithms, and regulators embracing this change through new collaborations. This perspective summarizes insights, recent developments, and recommendations for infusing actionable computational evidence into clinical development and health care from academy, biotechnology industry, nonprofit foundations, regulators, and technology corporations. Analysis and learning from publicly available biomedical and clinical trial data sets, real-world evidence from sensors, and health records by machine-learning architectures are discussed.
The need for new medical treatments and drugs has never been greater. But before pharmaceutical companies can go to market with a breakthrough drug, they need to ensure safety and efficacy through clinical trials. While this process is essential, it's also slow, expensive and unpredictable. Pharma R&D teams are solving this problem by leveraging the power of artificial intelligence (AI) in clinical trials to save time and money.
According to the US Food and Drug Administration (FDA)[8], approximately 33 percent of drugs move from Phase II to III, while around 25 to 30 percent move from Phase III to the next phase

Drug development is already a difficult endeavour, with the vast majority of R&D efforts failing to produce a market-worthy product. Even reaching the clinical trial phase offers no guarantees, as only 12% of such drugs receive U.S. Food and Drug Administration approval)[2]. Pharma companies need tools like AI that can reliably improve this percentage without jeopardizing safety. With the power of AI, companies can rapidly digitize clinical trial processes so they can complete studies faster. That means life-saving medicines and treatments can get to patients more quickly—and life sciences companies can gain a competitive edge.
"The big delay areas are always patient recruitment, site start-up, querying, data review, and data cleaning," explains Scott Clark, chief commercial officer a Taimei

Keywords: Clinical trials; design; study; Machine Learning; Artificial intelligence; Machine learning; Drug development; Precision medicine; monitoring; Predictive modeling.

ISSN: 0030-5324
UGC CARE Group 1

INTRODUCTION

Applying artificial intelligence (AI) and machine learning (ML) to clinical trials can significantly enhance various aspects of the process, from predicting outcomes to identifying trends and streamlining data analysis. Here's how these technologies can be utilized effectively. Clinical trials are research studies that test new treatments or interventions in people. They typically progress through several phases, each designed to answer different questions about the treatment's safety and efficacy. Here's a brief overview of each phase:

Phase 0 (Exploratory IND Studies):
This phase involves gathering preliminary data on how a drug behaves in the body with a very small number of healthy volunteers or patients. It focuses on pharmacokinetics (how the drug is absorbed, distributed, metabolized, and excreted) and pharmacodynamics (the drug's effects on the body). The duration is usually short, often just a few weeks.

Phase I (Safety and Dosage):
The main purpose of this phase is to evaluate the safety of a new drug or treatment and determine a safe dosage range. It typically involves 20 to 100 healthy volunteers or patients. The focus is on assessing side effects, identifying a safe dosage, and understanding how the drug behaves in the body. This phase generally lasts several months.

Phase II (Efficacy and Side Effects):
In this stage, the drug's effectiveness is assessed and its safety is further evaluated. Participants usually include 100 to 300 patients who have the condition the drug is intended to treat. The key focus is on evaluating how well the drug works and monitoring for adverse effects. The duration may range from several months to a couple of years.

Phase III (Confirmation and Comparison):
This phase confirms the drug's effectiveness, monitors side effects, and compares it to existing treatments. It involves large groups of patients, often 1,000 to 3,000 or more, and is usually conducted across multiple centers to ensure representation of diverse populations. The focus is on gathering comprehensive data on both efficacy and safety. This phase can last for several years.

Phase IV (Post-Marketing Surveillance):
After a drug has been approved for public use, Phase IV studies continue to monitor its long-term effects and overall effectiveness in real-world settings. This stage includes a large population of patients using the drug in daily practice. The focus is on identifying rare or long-term side effects, assessing the drug's long-term impact, and evaluating its performance in broader populations. Unlike earlier phases, Phase IV is ongoing and continues for as long as the drug is on the market.
Each phase is crucial for ensuring that new treatments are both safe and effective before they become
Widely available

Journal
Of the
Oriental Institute
M.S. University of Baroda

ISSN: 0030-5324
UGC CARE Group 1

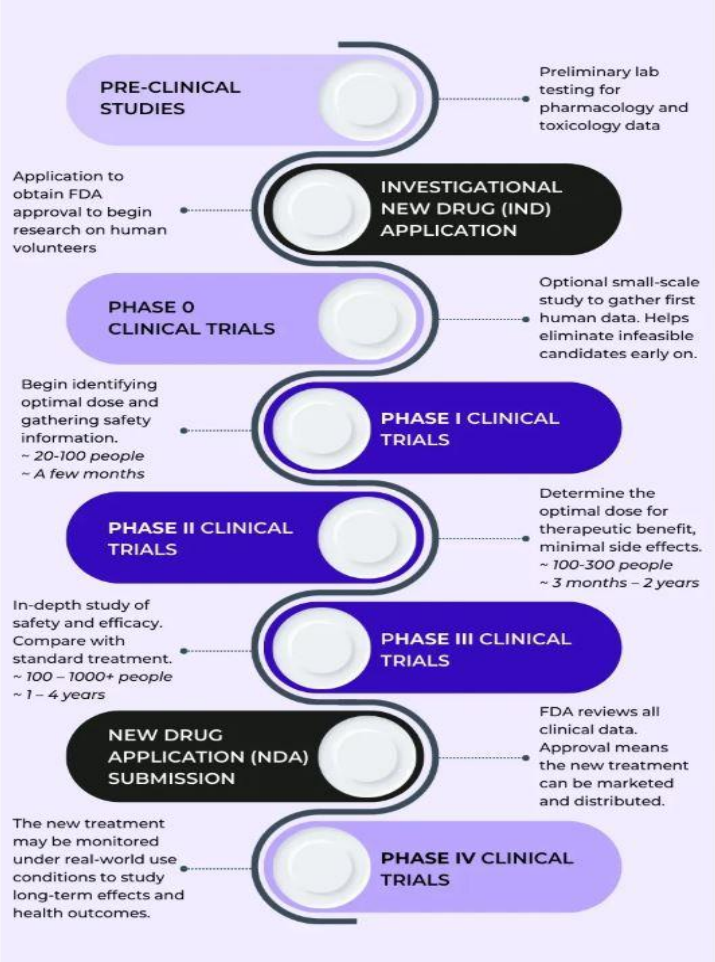

Fig: 1 Clinical trial phases

LITERATURE REVIEW

Clinical trials are crucial for evaluating the safety and efficacy of new treatments. However, the traditional process of clinical trials is often time-consuming and costly, with studies sometimes taking years to complete. Recent advancements in Machine Learning (ML) and Artificial Intelligence (AI) offer promising solutions to enhance the efficiency and speed of clinical trials. This literature review explores the current state of research on leveraging ML and AI algorithms to accelerate various phases of clinical trials, including design, recruitment, monitoring, and data analysis.

ML and AI algorithms have demonstrated significant potential in optimizing clinical trial design. For instance, adaptive trial designs, which use real-time data to modify the study protocol, have been enhanced by AI techniques. Bowalekar SK discuss how AI algorithms can dynamically adjust trial parameters based on ongoing data, reducing the time required to reach conclusive results. Similarly, algorithms that predict patient responses to treatments can help in designing more effective trial protocols.

A study by Liu J [5]et al. (2023) highlights how machine learning models can sift through vast datasets to match patients with trials, improving recruitment rates and reducing the time to enroll participants. Recruitment is a critical and often challenging phase of clinical trials. Traditional methods of recruitment are frequently inefficient and slow. AI-driven tools are transforming this process by analyzing electronic health records (EHRs) and other patient data to identify suitable candidates more quickly.

A study by Huhn S [4]et al. (2022) highlights Real-time monitoring of patients during clinical trials is essential for ensuring safety and efficacy. AI algorithms are increasingly used for continuous monitoring through wearable devices and digital health tools. For instance, AI systems can analyze data from wearable sensors to detect adverse events or deviations from expected responses. This enables quicker responses to potential issues, thereby safeguarding participants and maintaining trial integrity.

Data analysis is one of the most time-consuming aspects of clinical trials. ML algorithms, particularly those involving natural language processing (NLP) and predictive analytics, are proving effective in analyzing complex datasets. According to a study by Teodoro D[6] et al. (2025), ML techniques can rapidly process and interpret large volumes of clinical data, identifying patterns and insights that might be missed by traditional methods. This accelerates the decision-making process and can lead to faster trial completion.

Several case studies illustrate the practical applications of ML and AI in clinical trials. For example, IBM Watson for Clinical Trials uses AI to match patients with trials and optimize study designs. Another case study by Jin S[8] et al. (2025) demonstrated how ML algorithms were used to predict patient outcomes and streamline trial protocols, leading to a reduction in trial duration by 25%.

Despite the promising advancements, there are challenges associated with integrating ML and AI into clinical trials. Issues such as data privacy, algorithmic bias, and the need for high-quality data are understand and trust their decisions Hulstaert [7] et al. (2024). Addressing these challenges is crucial for the successful implementation of AI-driven solutions in clinical trials.

Tam A[12] et al. (2021) has represented Medical image processing benefits greatly from the application of artificial intelligence and machine learning, particularly deep learning techniques. These techniques enable the automation of previously manual tasks such as organ, tissue, and structure segmentation. Additionally, by minimizing human engagement and lowering human-induced variability that might bias the results, the use of automatic AI tools reduces the number of patients required to achieve higher statistical power, hence speeding up the time it takes for new compounds to reach the market.

Zhang X has presented[10] a historical analysis of the development highlighting key events that shaped the advancement of scientific and computational developments that now form AI/ML Qian L[9] et al. (2024) is confident about the AI/ML applications to grow substantially in near future. Automation, machine learning, and decision-making in continuous process control have greatly benefited modern manufacturing, which is where AI/ML in manufacturing takes traditional models. These advancements could not have been made without human engagement. Even while AI and ML increase productivity, this won't have an effect on the labor market because people are still required to develop the algorithms, monitor their use, and ultimately assess their results.

FDA is also learning from these models and making necessary changes in the guidelines and policies to support the quicker drug process and development

ISSN: 0030-5324
UGC CARE Group 1

have implemented a continuous learning system (CLS) based on deep learning and optimization and ensemble techniques, and carried out a retrospective data simulated prospective study using ultrasound images of breast masses for accurate diagnoses. This method can also be used for other organs. overall, their research may help advance the use of AI diagnosis in precision medicine.
Kolluri, S. and team[13] has aimed to clarify key concepts, present use-cases and finally offer insights and a balanced view on the best way to apply AI/ML methods in R&D.

Gupta, R., Srivastava, and D. Sahu[14] have given a summary of how artificial intelligence (AI) and traditional chemistry can be used to improve the drug discovery process, as well as how AI can be used to enhance the current drug discovery process. Next, they go over the various uses of AI in drug design and discovery, including polypharmacology, drug-target interactions, drug toxicity, drug repositioning, drug dosage effectiveness and efficacy, drug toxicity, drug toxicity, drug release and monitoring, and primary and secondary screening.

Muller, C., Rabal, O., Diaz Gonzalez's research work[15] provides the state of the art of AI methods applied to drug discovery with a focus on structure and ligand-based virtual screening, de novo design, chemical reactions, library design and high-throughput analysis, drug repurposing and drug sensitivity, de novo design, chemical reactions and synthetic accessibility, ADMET, and quantum mechanics.
Selvaraj, C., Chandra, I. & Singh, S.K[16] has discussed backend of AI and ML methods in supporting the drug design with machine learning, along with AI and ML challenge and opportunity for the pharmaceutical industry.

They also stated that the pharmaceutical industry can benefit greatly from the application of AI's ML and DL techniques to comprehend lead molecule chemical structures and activity relationships from a variety of pharmaceutical data sets. Due to AI-based technologies' improved data mining capabilities, automated medication development has started to utilize them lately.

MATERIALS & METHODS

A. Predicting Outcomes

1) Predictive Modeling

- Risk Stratification: Develop models to predict patient responses to treatments based on historical data, genetic information, and other biomarkers.
- Outcome Prediction: Use ML algorithms to forecast trial outcomes based on early results.
- Linear Regression: Predict continuous outcomes like disease progression or treatment response.
- Logistic Regression: Used for binary outcomes, such as predicting whether a patient will experience a particular side effect.

2) Personalized Medicine

- Patient Stratification: Apply AI to analyze patient data and identify subpopulations more likely to respond to specific treatments.

ISSN: 0030-5324
UGC CARE Group 1

- Support Vector Machines (SVM): Classify patients into groups based on treatment response likelihood.
- Random Forests: Handle large datasets and provide feature importance for outcome prediction.
- Neural Networks: Deep Learning models (CNNs for image data, RNNs for sequential data) to predict patient outcomes.

B. Identifying Trends

1) Pattern Recognition

- Data Mining: Use AI to uncover hidden patterns in large datasets.
- Longitudinal Analysis: Apply ML to track changes over time.
- LSTM Networks: Specialized RNNs for sequential data like patient vitals.

2) Early Signal Detection

- Safety Monitoring: AI monitors adverse event reports to detect safety signals.
- K-Means Clustering: Group patients based on similarities.
- Hierarchical Clustering: Detect natural groupings in patient responses.

3) Dimensionality Reduction

- Principal Component Analysis (PCA): Reduce features while retaining significant variability.

C. Streamlining Data Analysis

1) Automation of Data Processing

- Data Cleaning: Automate cleaning processes with AI.
- Integration of Diverse Sources: ML integrates EHRs, wearables, and lab results.
- Named Entity Recognition (NER): Extract key information from unstructured data.
- Topic Modeling (LDA): Identify key themes in text data.

2) Advanced Analytical Techniques

- Natural Language Processing (NLP): Analyze unstructured text data.
- Predictive Analytics: Simulate trial scenarios to optimize design.
- Entity Resolution: Merge datasets from different sources.

3) Anomaly Detection

- Isolation Forest: Identify outliers in trial data.
- Autoencoders: Neural networks for anomaly detection in high-dimensional data.

D. Enhancing Data Quality and Reliability

1) Anomaly Detection

- Outlier Detection: Detect anomalies or inconsistencies to identify errors or fraud.

2) Quality Assurance

- Automated Monitoring: AI monitors data quality in real-time.

ISSN: 0030-5324
UGC CARE Group 1

E. Optimizing Trial Design and Execution

1) Trial Simulation

- In Silico Trials: Simulate trials virtually to predict outcomes.
- Monte Carlo Simulation: Assess risk with simulated trial scenarios.
- Genetic Algorithms: Optimize trial parameters via evolutionary processes.

2) Resource Allocation

- Efficient Design: Optimize sample size, dosing, and endpoints.
- Bayesian Optimization: Dynamically adjust trial parameters based on interim results.

F. Facilitating Real-Time Decision Making

1) Decision Support Systems

- AI-Driven Insights: Provide real-time recommendations.
- CART: Decision trees for complex datasets.

2) Dynamic Adjustments

- Adaptive Trials: Modify protocols in real-time based on outcomes.
- Gradient Boosting Machines (GBM): Combine models for accuracy.
- XGBoost: Scalable gradient boosting for predictive tasks.

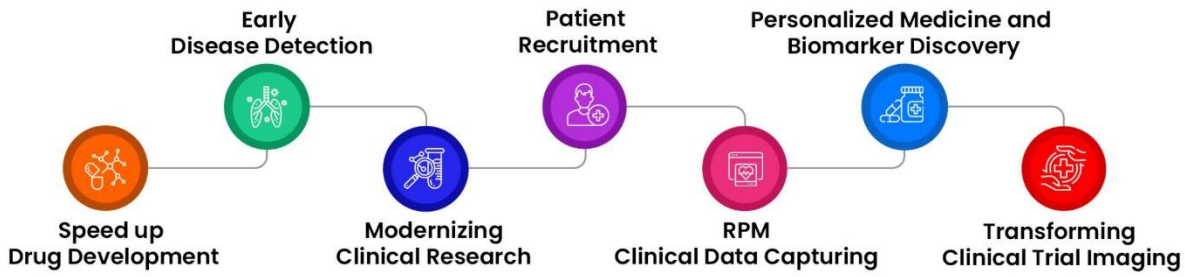

Fig: 2 Use of AI/ML in different phases of clinical trial

CONCLUSION

In conclusion, the application of AI and machine learning to drug development is an exciting and rapidly evolving field with great promise to speed up the identification of new medications and give patients access to the most advanced care.

Category	Key Results
A. Predicting Outcomes	More accurate patient response prediction, early outcome forecasting, and improved personalized treatment.
B. Identifying Trends	Detects hidden patterns, tracks patient progress, and enables early safety signal detection.
C. Streamlining Data Analysis	Automates cleaning, integrates diverse data, and accelerates analysis with anomaly detection.
D. Enhancing Data Quality	Ensures reliable datasets through real-time monitoring and error/fraud detection.
E. Optimizing Trial Design	Enables virtual simulations, optimizes dose/sample size, and reduces trial cost and time.
F. Real-Time Decision Making	Supports adaptive trials, provides AI-driven recommendations, and enables rapid adjustments.

Table:1 AI/ML Benefits in Clinical Trials

There are still certain problems and limitations that need to be overcome before machine learning may be completely utilized in drug discovery. These include the need for additional high-quality data, more cooperation and data exchange, transparency and visibility in the use of machine learning models, and moral and legal issues. There is a great deal of scope for further progress in the field of machine learning in drug discovery. If we work together to find solutions to these problems, we can leverage the potential of machine learning to expedite the drug development process and produce more effective medications for patients in need.

Declaration by Authors
Ethical Approval: Approved
Acknowledgement: None
Source of Funding: None
Conflict of Interest: No conflicts of interest declared.

REFERENCES

1. Bowalekar SK. Adaptive designs in clinical trials. *Perspect Clin Res*. 2011 Jan–Mar;2(1):23–27. doi:10.4103/2229-3485.76286.
2. Ramachandran A. AI-Driven Clinical Trial Optimization: Reducing Patient Recruitment Time and Enhancing Success Rates [Internet]. 2024 Oct [cited 2025 Sep 05]. Available from: https://www.researchgate.net/publication/385418862_AI-Driven_Clinical_Trial_Optimization_Reducing_Patient_Recruitment_Time_and_Enhancing_Success_Rates

3. Javaid M, Haleem A, Singh RP. Health informatics to enhance the healthcare industry's culture: an extensive analysis of its features, contributions, applications and limitations. *Informatics and Health*. 2024;1(1):123–48. doi: 10.1016/j.infoh.2024.05.001.
4. Huhn S, Axt M, Gunga H-C, Maggioni MA, Munga S, Obor D, et al. The impact of wearable technologies in health research: scoping review. JMIR Mhealth Uhealth. 2022 Jan 25;10(1):e34384. doi: 10.2196/34384.Shah, P., Kendall, F., Khozin, S., *et al.* (2019). *Artificial intelligence and machine learning in clinical development: A translational perspective. Nature Digital Medicine, 2*, 69. https://doi.org/10.1038/s41746-019-0148-3
5. Liu J, Allen PJ, Benz L, Blickstein D, Okidi E, Shi X. A machine learning approach for recruitment prediction in clinical trial design. arXiv [Preprint]. 2021. doi:10.48550/arXiv.2111.07407.
6. Teodoro D, Naderi N, Yazdani A, Zhang B, Bornet A, et al. A scoping review of artificial intelligence applications in clinical trial risk assessment. npj Digit Med. 2025;8:486. doi:10.1038/s41746-025-01886-7.
7. Hulstaert L, Twick I, Sarsour K, Verstraete H. Enhancing site selection strategies in clinical trial recruitment using real-world data modeling. PLoS One. 2024;19(3):e0300109. doi:10.1371/journal.pone.0300109.
8. Jin S, Chen L, Ding H, Wang M, Yu L. Can artificial intelligence predict clinical trial outcomes? arXiv [Preprint]. 2025. doi:10.48550/arXiv.2411.17595.
9. Qian L, Lu X, Haris P, Zhu J, Li S, Yang Y. A survey of artificial intelligence methods for clinical trial outcome prediction. ChemRxiv [Preprint]. 2024. [No DOI available]
10. Zhang X, Xiao C, Glass LM, Sun J. DeepEnroll: patient-trial matching with deep embedding and entailment prediction. arXiv [Preprint]. 2020.
11. Biswal S, Xiao C, Glass LM, Milkovits E, Sun J. Doctor2Vec: dynamic doctor representation learning for clinical trial recruitment. arXiv [Preprint]. 2019.
12. Tam A, Laurent C, Gauthier S, Dansereau C. Prediction of cognitive decline for enrichment of Alzheimer's disease clinical trials. arXiv [Preprint]. 2021.
13. Kolluri S, Lin J, Liu R, et al. Machine Learning and Artificial Intelligence in Pharmaceutical Research and Development: a Review. AAPS J. 2022;24:19. doi:10.1208/s12248-021-00644-3. PubMed
14. 14. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Molecular Diversity. 2021;25:1315–1360. doi:10.1007/s11030-021-10217-3. PubMedDigital Diagnostics
15. 15. Müller C, Rabal O, Díaz Gonzalez C. Artificial Intelligence, Machine Learning, and Deep Learning in Real-Life Drug Design Cases. In: Heifetz A, editor. Artificial Intelligence in Drug Design. Methods in Molecular Biology, vol 2390. New York, NY: Humana; 2022. p. 383–407. doi:10.1007/978-1-0716-1787-8_16. Europe PMCSpringer Nature Experiments
16. Selvaraj C, Chandra I, Singh SK. Artificial intelligence and machine learning approaches for drug design: challenges and opportunities for the pharmaceutical industries. Molecular Diversity. 2022;26:1893–1913. doi:10.1007/s11030-021-10326-z.