Review Article

The Jour Multi Rese, 6(2), 2026, 10-16

The Journal of Multidisciplinary Research (TJMDR)

Content Available www.saap.org.in ISSN: 2583-0317

AI DRIVEN CLINICAL TRIALS: REDEFINING MEDICAL INNOVATION AND RESEARCH EFFICIENCY

THAGARAKUNTA BHARANESWAR*, NADAKUDURU ASHOK, BUDAGALA GAYATHRI, CHANDU BABU RAO.

Priyadarshini Institute of Pharmaceutical Education and Research, 5th Mile, Pulladigunta, Vatticherukuru (M), Guntur-522017, Andhra Pradesh, India.

Article History: Received: 11 Apr 2026 Revised: 22 Apr 2026 Accepted: 20 June 2026

Abstract: Artificial Intelligence (AI) has made a noticeable impact on the pharmaceutical industry, transforming several key areas of its operations. Artificial intelligence (AI) is turning the field of clinical trial design, execution, and analysis around by undertaking long-standing issues of high expenses, extended periods, low success rates, and inefficiencies in patient and data number recruitments. AI has been used in clinical trials in protocol optimization, feasibility, analysis of trial outcomes, intelligent patient identification, real time monitoring of data and improved patient engagement. The recent integration of technology providers with pharmaceutical companies has increased the step of practical application of AI in clinical research. The collaboration of Pfizer with IBM Watson to recruit patients to clinical trials, Roche with Google Cloud to develop a digital biomarker, AstraZeneca with Benevolent AI to develop a drug using AI, and others reflect the increasing role of AI in the clinical trial ecosystems. The future of AI application in clinical trials will have to rely on further technological improvement, regulatory modifications, and interdisciplinary cooperation. Moreover AI Can enhance the trial efficiency and cost effective clinical research.

Keywords: *Artificial Intelligence, Drug Discovery, Drug development, Clinical trials, Machine Learning, Personalized Medicine.*

***Corresponding Author**
Thagarakunta Bharaneswar

Produced and Published by
South Asian Academic Publications

INTRODUCTION

In the pharmaceutical industry, the challenge to come up with new therapeutics is increasing with innovation forcing new financial and scientific resources into the sector. Recent reports indicate that worldwide Pharma R&D expenditure is estimated to be around $250 billion annually with less than 14% of drug candidates making it to approvals allowing average development costs to be well over $2.6 billion per approved drug [1]. This disparity between input and output demonstrates that there is an immediate necessity to more effectively; data-centric approaches to increase the productivity of research efforts and reduce churn in the drug pipeline.

Artificial intelligence (AI) has become a key technological component that is transforming various sectors around the world in recent years. Once at the core of big tech companies, AI technologies are turning into the most active influencer of the biomedical field and healthcare overall, improving disease detection, aiding precision medicine, improving patient outcomes, and decreasing healthcare spending [2]. The machine learning (ML), deep learning (DL), and artificial neural networks (ANNs) subfields of AI are now transforming the pattern of clinical and pharmaceutical sciences. These computational platforms simulate human intelligence in order to simplify the drug discovery process, minimize the errors during drug development, predict therapeutic targets, and assist the design of the tailored treatment regimen in an even more reliable and efficient manner. With this growing technological environment, it is necessary to conduct a deep evaluation of the potential of AI with the view of informing future pharmaceutical and clinical practice [3].

1. BASIC CONCEPTS OF AI USED IN HEALTH CARE

The world has seen a significant revolution in information technology (IT) over the past fifty years, which has led to the creation and storage of large amount of data in various domains, including technology, and has provided researchers with a variety of services and goods [4]. Big data now has unmatched potential in terms of enhancing medical outcomes and public health [5]. Machine learning (ML) is one of the most widely utilized approach in artificial intelligence (AI). Machine Learning (ML), a major branch of Artificial Intelligence, is broadly classified into three categories: Supervised learning, Unsupervised learning, and Reinforcement learning. In supervised learning, the model is trained using labelled datasets where both input and output data are known; it learns the relationship between them and is used to make predictions such as disease diagnosis from patient records or classifying medical images as normal or abnormal. Unsupervised learning deals with unlabelled data and identifies hidden patterns or groupings within the dataset; this is useful in healthcare for tasks like grouping patients based on similar symptoms, identifying unknown disease types, and detecting abnormals in medical reports. The third type, reinforcement learning, works on a trial-and-error basis where the algorithm interacts with the environment and receives feedback in the form of rewards or penalties, enabling it to learn the best possible actions over time; this approach is applied in areas such as robotic surgery, treatment optimization, and adaptive insulin dosing systems. Suitable Vector Machine used in clinical trials to classify patients, predict eligibility, and detects adverse drug reaction; Artificial Neural Networks are applied to predict clinical outcomes, treatment response, and PK/PD relationships. Convolutional Neural Networks are used for automated analysis of medical images and pathology data in trials. Recurrent Neural Networks are used to analyze time-series clinical data such as disease progression, vital signs, and longitudinal trial outcomes.

2. AI IN EARLY STAGE AND PRE-CLINICAL DRUG DEVELOPMENT

In contrast to traditional drug discovery, which may require more than 10 years, and where laboratory and animal assays play a major role, AI can process large volumes of data in a short time, forecast molecular behaviour, and discover safety concerns early. This simplifies the process of discovery, minimizes expenditure and usage of animals and enhances the probability of identifying successful drug candidates to undergo clinical trials.

2.1 Target Identification and Validation

This is the process of finding a gene or a protein that is related to the disease and its potential as a therapeutic agent. Benevolent AI, Insilco Medicine, and Schrodinger are examples of AI platforms that helps in the analysis of biological data, target relevance predictions, and off-target and safety risk assessment. This combination of AI enhances efficiency and lowers drug development failure.

2.2 Molecule Design

Machine learning and deep learning models are used in AI to examine chemical and biological data in molecule design, which allows rapid discovery, optimization and creation of novel compounds. Methods like structure-based drug design, QSAR modelling methods, and de novo molecular generation improve the specificity of the target and efficiency of the pipeline.

2.3 Toxicology and Safety Prediction

In toxicology and safety prediction, AI predicts ADME properties and possible toxicities earlier on during development, which enables high-risk compounds to be dropped earlier. The ongoing process of learning using new information enhances better prediction and promotes real-time decision-making and it also complies with ethical norms of research.

2.4 Simulation Models

In-silico testing of drugs in the population and dosing situations is possible with AI-based simulation models like organ on chip models, In-silico experiments, PK/PD models, virtual screening, and digital twins. These devices reduce the workload of the experiment and equip candidates with clinical trials. Generally, AI develops an integrated efficient innovative preclinical research ecosystem that boosts productivity, safety, and translation success [8].

3. AI BASED IMPROVEMENTS IN PLANNING CLINICAL TRIALS

3.1 Protocol Design

The protocol optimisation provided by AI is a paradigm shift in the design of trials, which relies on intuition, to design based on evidence. State-of-the-art machine learning (ML) models are used to process large amounts of historical trial data and find patterns that can be used to guide the optimal study parameters [6]. The process of natural language processing (NLP) algorithms systematically chews on what the thousands of previous study protocols, regulatory sub missions, and published literature have to say, to suggest evidence-based inclusion/ exclusion criteria [7]. These AI technologies are already shown to be much more effective than traditional statistical technologies: whereas traditional regression-based methods handle fewer than 65% of the protocol optimisation accuracy, machine learning algorithms can consider thousands of variables at once, withbest results of 80% success. Deep learning models are capable of making predictions related to the feasibility of the protocols by considering several variables at a time, such as population traits of the target, spatial distribution, seasonal variations, and competitive environment. These predictive models achieve accuracy rates exceeding 80 % in estimating enrolment success, significantly out performing traditional feasibility assessments [9]. A notable example is Google's retinal imaging system that achieved high accuracy in predicting patient gender from retinal photographs (despite male and female retinas being anatomically identical), emphasizing the risk of algorithms identifying clinically irrelevant but statistically significant patterns [10].

3.2 Feasibility Assessment

The algorithms of machine learning transform the site selection process by combining various data streams and estimating the performance and recruitment capacity of the investigator. These models are used to analyze demographic data, occurrence of diseases, healthcare infrastructure, experience of investigators and the performance metrics of sites in the past to prioritize the potential sites of study [11]. More recent geospatial analysis software includes the socioeconomic factors, transport accessibility, and competing trial activity to maximise site selection strategies. The AI tools can predict the enrolment rates on a site level which allows the timeline to be and time losses, improve the feasibility of trials, and improve the chances of successful drug development.

more accurately forecasted and the resources to be shared, minimising the chances of study delays and cost overruns.

3.3 Early Decision Support

Predictive analytics and AI tools are applied during the initial planning phases of a clinical trial assess the need to continue or alter a study. AI uses big biomedical data, which includes preclinical outcomes, outcomes of past trials, demographics of the patients, and disease-specific data, to predict the likelihood of the trial success. It helps in detecting risks, narrowing of inclusion and exclusion criteria, optimal dosing schedules and better endpoint selection and anticipates possible delays such as ineffective recruitment or safety issues. On the whole, early decision support through AI will improve the evidence-based decision making, minimize financial

Table 01: Performance of AI Applications in Planning Clinical Trials compared to normal traditional methods.

Application	Description	Traditional Method	AI Impact (%)	Advantage
Protocol Design	AI Analyses (NLP+ML) Previous data and suggest optimized dosage and parameters.	65% Prediction accuracy	80% of the design process can be fastened by AI	Reduces protocol diversions, improves clarity and efficiency.
Feasibility Assessment	Predictive algorithms(ML) use historical and present data to give recruitment rate, eligible patients list and trail costs	8-12 Weeks Planning time	50-70% increase in accurate feasibility evaluation	Prevents poor site selection, reduce trial delays and improves planning.
Early Decision Support	AI (DL) predicts all the data and tells whether to proceed or not.	55% decision accuracy	75% of process is done by AI tools	Saves time and money by avoiding low probability trials & improves success rate.

4. SMART PATIENT IDENTIFICATION AND ENROLLMENT USING AI

Patient recruitment represents the most critical bottleneck in clinical trial conduct, with AI offering transformative solutions through automated participant identification and matching systems. Advanced Natural language processing (NLP) algorithms scan electronic health records (EHRs), clinical notes, and laboratory results to identify potentially eligible participants with unprecedented efficiency and accuracy [12]. IBM Watson for clinical trial matching showed 78% accuracy and reduced screening time by 78%, highlighting AI's potential benefits. However, real-world implementation faced major challenges. Integration with hospital EMR systems cost $300,000–$600,000 per case, limiting economic feasibility. This made adoption difficult even in high-budget oncology trials. Machine learning models can also predict patient enrollment success using historical, demographic, and clinical data. [13]. AI systems significant potential to address

Longstanding disparities in clinical trial participation, mainly in terms of racial, ethnic, socioeconomic representation. Natural language processing tools can analyse social determinants of health, transportation barriers, and cultural factors that influence trial participation, informing the development of more inclusive recruitment strategies. After all the stages i.e. participant identification and recruitment disparities, virtual screening is done to continue the clinical trial process. AI-powered virtual screening platforms enable remote participant evaluation, reducing geographical barriers to trial participation. Computer vision algorithms can analyse medical images, while NLP tools process patient-reported outcomes and digital health data to conduct preliminary eligibility assessments. Intelligent chatbots and conversational AI systems facilitate remote consent processes, providing personalised information delivery and addressing participant questions in real-time [14]. When coming to the risk factors the use of AI in patient recruitment is risky and needs proper management. Less risky tools like estimation of the cohort size are not that effects and can be confirmed statistically. Risky applications such as optimization of recruitment strategies can be impactful on timelines and are tested in controlled settings. Patient eligibility decisions and other high-risk apps need rigorous reviews by the physicians and investigators to ensure the safety of the patients.

5. DIGITAL DATA CAPTURE AND MONITORING THROUGH AI TOOLS

The traditional clinical trials were very manual with data being keyed in through paper, case report forms and monthly site visits and that resulted in error in data, time and incomplete reporting. AI powered digital applications solve these short comings by allowing the collection of all the data collected in real time and precisely and continuously during the trial period.

5.1 Electronic Data Capture (EDC) Systems

Electronic Data Capture systems are computerized systems that are employed to receive, store and handle clinical trials information as an alternative to paper-based report forms. EDC systems are also intelligent when integrated with AI and can guarantee quality data throughout the entry into the system. Depending on predetermined clinical and protocol rules, AI algorithms automatically check data by determining missing values, inconsistencies, duplicates, and out-of-range values. Besides that, machine learning models facilitate risk-based monitoring, whereby important data fields and high-risk offices are put on the priority list to review. This minimizes the manual monitoring burden and ensures regulation is done. AI-driven EDC systems produce real-time dashboards that enable sponsors and investigators to monitor the recruitment, protocol compliance, and safety indicators effectively and more promptly and make informed decisions.

5.2 Wearables and Sensor based devices

A sensor-based device and Wearable's are important in obtaining continuous and objective data throughout the clinical trial. Smart watches, fitness bands, biosensors, and implantable sensors are the types of devices that are used to record real-time physiological data such as heart rate, blood pressure, physical activity, sleep patterns, and glucose levels. AI is used to process the high amounts of information produced by these devices to recognize trends, identify irregularities, and produce valuable clinical insights. In comparison to the conventional episodic monitoring performed at the time of the visit to the site, the data obtained by using a wearable will be a more realistic view of the health of a patient in the real-life conditions. This enhances endpoint sensitivity, minimizes reporting bias, and promotes decentralized models of clinical trials.

5.3 Remote Patient Monitoring

Remote patient monitoring involves AI-based mobile apps and telehealth devices to monitor patient health and compliance with treatment beyond clinical locations. Patients will be able to share their symptoms, medication history, and unpleasant experiences via online tools and AI machines will constantly process this information to give timely alerts. This will reduce the number of visits to the hospital, increase the level of convenience to the patient, and increase the recruitment and retention rates. The AI-powered notifications inform the investigators about the possible safety concerns or non-adherence promptly to intervene accordingly. Remote monitoring is particularly useful in longer studies and trials of long-lasting trials or elderly populations or geographically far populations.

5.4 Automated Data Cleaning and Integration

AI-powered data cleaning tools address one of the most time- consuming aspects of clinical trial conduct, automatically identifying and correcting data inconsistencies, missing values, and entry errors. Natural language processing algorithms standardise free-text entries, whilse machine learning models predict missing data points based on patient characteristics and study context [15]. Traditional data cleaning processes require 60-80 hrs of biostatistician time per 100 patient datasets, with manual review and query resolution extending timelines by 4-6 weeks. AI-powered systems reduce this to 12-16 hrs of oversight time with automated processing completed within 24-48 hrs [16]. These automated systems reduce data cleaning time by 60-80 % whilst improving data quality and consistency across study sites. Advanced imputation algorithms, including deep learning, generative adversarial networks, and recurrent neural networks, can handle complex matters.

6. PREDICTIVE MODELS FOR UNDERSTANDING TRIAL OUTCOMES

6.1 Regression Models

Regression models are clinical outcome predictions that quantify the relationship between treatment related variables (dose, exposure, demographics) and responses, and as such are directly applicable to dose response assessment, efficacy prediction, and toxicity. For example, Regression predictors are employed in a Phase II antihypertensive trial to estimate the reduction in blood pressure levels at varying dosage levels and an appropriate dose is selected that will produce optimal efficacy with the least adverse effects. Figure 05(A) shows linear regression, a straight line shows outcome vs predictor, Figure 05(B) shows Logistic regression, a S-Shaped curve shows Probability of response.

6.2 Survival Analysis Models

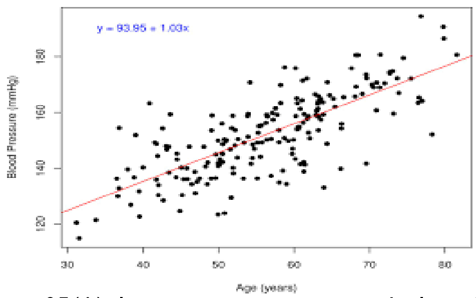
The models of survival analysis are used to make predictions regarding time-to-event outcomes, e.g. disease progression, relapses, or death, and are commonly used to calculate progression-free survival, overall survival and time to adverse event in clinical trials. For example, Survival analysis can be applied in oncology trials to predict the duration of a patient without progressing following a novel anticancer therapy to enable researchers to measure long-term clinical benefit at interim periods. Figure 05(C) shows the graphical data which X-axis shows time and Y-axis shows Probability of survival, step down curves for treatment group and control/placebo group.

6.3 Machine Learning Models

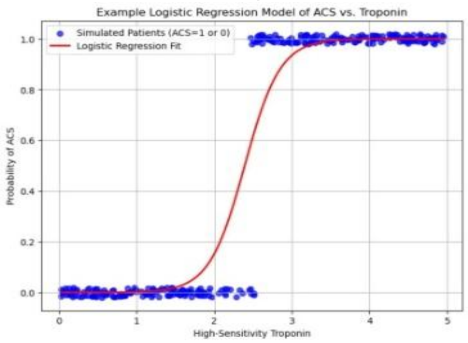
Machine learning models are models that predict the outcome of a trail based on the learning of a complex and non-linear pattern using large datasets and can be used to predict treatment response, risk stratifying patients, predicting the probability of trial success, and so forth. For example, Machine learning algorithms are based on analysing genomic data and clinical data to forecast the patients who are most likely to respond to specific cancer treatments, enhancing the selection of patients and decreasing the rate of trial failure. Figure 05(D) shows model training.

6.4 Bayesian Models

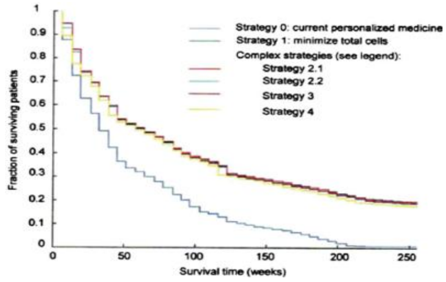
Bayesian models are probability-based predictions models that are capable of continuously updating with the appearance of new data making them appropriate to adaptive trial designs, interim analysis, and early decision-making. For example, Bayesian models are used in adaptive clinical trials to make decisions regarding early termination of ineffective arms of treatment, or to increase promising arms of treatment based on the accumulating evidence of efficacy and safety. Figure 05(E) shows the graph data which X-axis shows treatment effect and Y-axis shows Probability. Contains three curves Prior distribution: before trial, likelihood: info about trail data, Posterior distribution: updated belief after seeing data.



Figurer 05(A): linear regression, a straight line shows outcome vs predictor



Figurer 05(B): Logistic Regression, S-Shaped Curve shows Probability of Response



Figurer 05(C): Step down Curves for Treatment Group and Control Group

Figurer 05(D): Three Curves for Prior Distribution likelihood and Posterior Distribution

7. REAL WORLD IMPLEMENTATIONS OF AI

7.1 Twist Bioscience and Oracle Partnership

Twist Bioscience partnered with Oracle Cloud Infrastructure (OCI) to enhance its data management and integration capabilities, focusing on protein structure prediction using advanced AI models like AlphaFold2, ParaFold, and ESMFold. Oracle's high-performance GPU instances and cloud storage solutions enabled Twist to significantly improve the efficiency and speed of its AI models, achieving up to an 82% increase in prediction speed and notable cost savings. This collaboration also set the stage for future innovations in DNA-based digital data storage, demonstrating Twist's potential to revolutionize data archiving on an unprecedented scale.

7.2 Pfizer – IBM Watson

Pfizer partnered with IBM Watson to use artificial intelligence in patient recruitment and oncology trial matching. IBM Watson is a machine-learning platform and natural language-processing (NLP) that analyzes electronic health records, pathology reports and clinical notes to identify potential cancer patients suitable to participate in a certain clinical trial. This cooperation has enabled the shortening of the recruitment time frames, enhancement of patient trial matching accuracy and one of the largest reasons of clinical trial delays-slow inefficient enrolment.

7.3 Novartis-Microsoft

Novartis collaborated with Microsoft to incorporate AI and cloud-based analytics in the design and execution of clinical trials. Novartis uses predictive models on historical clinical trial data to design and select sites and patient retention strategies using the Azure AI and data platform provided by Microsoft. This integration helps in adaptive trial design and real-time trial monitoring that in the long run will enhance efficiency and decision-making.

7.4 Johnson & Johnson-Medidata

J&J collaborated with Medidata to improve AI-related clinical data monitoring and risk management. The AI-based systems of Medidata allow cleaning of the data in real-time, centralized statistics surveillance, and the identification of protocol deviation or safety risk early. The collaboration enhances patient safety, regulatory compliance and data quality in global clinical trials.

8. ETHICAL, LEGAL AND REGULATORY ASPECTS

Informed consent, patient autonomy, data privacy and fairness are some of the ethical issues that AI in clinical trials would provoke. Data security problems and algorithmic bias on patient selection are some of the risks. The legal issues include ownership and responsibility of data and the problem does not have a universal regulation all over the world, although the EU has suggested the Artificial Intelligence Act (AIA) to provide the opportunity to use it safely and legally. [17]. It identifies some systems as high-risk and they have to be subject to mandatory conformity assessments, especially to medical equipment and equipment used in an in vitro diagnostic test in Articles 30 and 43 [18]. Transparency, responsibilities, testing environments, and penalties are also covered in the Act. The U.S. Food and Drug Administration, European Medicines Agency, and Indian Council of Medical Research are regulatory organizations that focus on the extensive validation of AI tools just like conventional clinical systems. Regulatory approval requires adherence to Good Clinical Practice (GCP), preservation of audit trails and human-in-the-loop decision-making. Unified ethical and regulatory bodies are needed to ensure responsible and safe AI use in the clinical trial.

9. CHALLENGES AND FUTURE PERSPECTIVES

Although it continues to gain usage, AI in clinical trials is facing significant challenges. The data are often of low-quality, not standardized, with the clinical trial data being partial, incomplete or biased. Ethical issues in regard to patient consent, data privacy and data protection regulations are still addresses. Black-box AI models are not transparent and explainable, and therefore it is hard to validate them and approve them regulatively. Other limitations are uncertainty in regulations, lack of AI skills in clinical researchers, high cost of implementation, and challenges in integrating into existing trial systems. AI has excellent prospects of enhancing efficiency, cost-effectiveness and trial success rates. Future trends include explainable AI models, data harmonization, federated learning, secure data-sharing, adaptive trial designs, digital twins and real-time decision-support systems. Academia, industry and regulators should collaborate to create standards of validation and ethics. The policymakers should be able to be fair, transparent and share the benefits fairly and equally and avoid digital disparities. International programs like GA4GH and the International Cancer Genome Consortium have shown the potential of international data sharing, although issues such as privacy of data, interoperability and geopolitical limitations still exist. These issues can be overcome with the help of federated learning and privacy-preserving computation [19, 20]. On the whole, AI should transform clinical trials and make them more predictive, patient-centered, and efficient.

10. CONCLUSION

The application of artificial intelligence (AI) can revolutionize clinical trials through better trial design, patient selection, data monitoring, and predictions of the outcome. It can hasten schedules, decrease costs and failures, decrease patient losses and improve the overall quality of research. Recruitment AI-based portals can boost enrolment by up to 65% whereas predictive analytics has the ability to predict trial outcomes at a rate of up to 85%. The presence of digital biomarkers and real-time analytics enhances the monitoring of safety and the efficiency of the trial.

Nevertheless, the obstacles present in the form of the high implementation cost (USD 350,000-700,000), poor data quality, algorithm bias, low transparency, regulatory complexity, and staffing constraints are still maturity, especially in the resource-constrained environments. Such obstacles can be eliminated with the help of various data sources, interpretable AI models, adaptable rules, employee education, and close cooperation between academia, industry, technology specialists, and regulators. A responsible and ethical development of AI is required in order to promote the improvement of drug development with regard to patient safety and scientific integrity.

11. AUTHOR CONTRIBUTIONS

All authors are contributed equally.

12. FINANCIAL SUPPORT

None

13. DECLARATION COMPETING INTEREST

The authors have no conflicts of interest to declare.

14. ACKNOWLEDGEMENTS

None

15. REFERENCES

1. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: new estimates of R&D costs. *J Health Econ*. 2016;47:20-33. doi:10.1016/j.jhealeco.2016.01.012.
2. Kenett RS. A perspective of artificial intelligence and clinical research [Internet]. SSRN; 2025 Jul 12 [cited 2026 Jun 22]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5361986

3. Chopra H, Shin DK, Munjal K, Dhama K, Emran TB. Revolutionizing clinical trials: the role of AI in accelerating medical breakthroughs. *Int J Surg*. 2023;109(12):4211-4220. doi:10.1097/JS9.0000000000000705.
4. Oliveira AL. Biotechnology, big data and artificial intelligence. *Biotechnol J*. 2019;14(8):1800613. doi:10.1002/biot.201800613.
5. Car J, Sheikh A, Wicks P, Williams MS. Beyond the hype of big data and artificial intelligence: building foundations for knowledge and wisdom. *BMC Med*. 2019;17(1):143. doi:10.1186/s12916-019-1382-x.
6. Awen BZ, Katakam P, Rao CB, Mohammed SA, Alokbe TO. Design and in-vitro evaluation of controlled release cephalexin subgingival films using natural biodegradable polymer. *Recent Res Sci Technol*. 2010;2(4).
7. Baluguri PK, Nama S, Chandu BR, Sakala B. LC/MS: an essential tool in drug development. *Int J Adv Pharm Anal*. 2012;1(2):24-37.
8. Solaiman B, Cohen IG, editors. *Research handbook on health, AI and the law* [Internet]. Cheltenham (UK): Edward Elgar Publishing; [cited 2026 Jun 22]. Available from: https://philpapers.org/rec/SOLRHO
9. Bae CY, Im Y, Lee J, Park CS, Kim M, Kwon H, Kim B, Park HR, Lee CK, Kim I, et al. Comparison of biological age prediction models using clinical biomarkers commonly measured in clinical practice settings: AI techniques vs. traditional statistical methods. *Front Anal Sci*. 2021;1:709589. doi:10.3389/frans.2021.709589.
10. Liu R, Rizzo S, Whipple S, Pal N, Pineda AL, Lu M, Arnieri B, Lu Y, Capra W, Copping R, et al. Evaluating eligibility criteria of oncology trials using real-world data and AI. *Nature*. 2021;592(7855):629-633. doi:10.1038/s41586-021-03430-5.
11. Betzler BK, Yang HH, Thakur S, Yu M, Quek TC, Da Soh Z, Lee G, Tham YC, Wong TY, Rim TH, et al. Gender prediction for a multiethnic population via deep learning across different retinal fundus photograph fields: retrospective cross-sectional study. *JMIR Med Inform*. 2021;9(8):e25165. doi:10.2196/25165.
12. Bohannon J, Balavarca Y, Basu S, Gerlach J, Guo W, Hergert D, et al. Artificial intelligence in clinical trials: a review of current applications and future opportunities. *Digit Biomark*. 2021;5(1):84-92.
13. Khader A, Ensan F. Learning to rank query expansion terms for COVID-19 scholarly search. *J Biomed Inform*. 2023;142:104386. doi:10.1016/j.jbi.2023.104386.
14. Hamer DM, Schoor P, Polak TB, Kapitan D. Improving patient pre-screening for clinical trials: assisting physicians with large language models [preprint]. arXiv. 2023 Apr 14. doi:10.48550/arXiv.2304.07396.
15. Hammond M, Ashford P, High J, Clark LV, Howard G, Jones M, Stirling S, West C, Norwich CTU Methodology Group. Designing e-consent protocols for pragmatic clinical trials: case studies from a UKCRC clinical trials unit. *Trials*. 2024;25(1):550. doi:10.1186/s13063-024-08386-1.
16. Côté PO, Nikanjam A, Ahmed N, Humeniuk D, Khomh F. Data cleaning and machine learning: a systematic literature review. *Autom Softw Eng*. 2024;31(2):54. doi:10.1007/s10515-024-00453-w.
17. Ebers M, Hoch VR, Rosenkranz F, Ruschemeier H, Steinrötter B. The European Commission's proposal for an Artificial Intelligence Act—a critical assessment by members of the Robotics and AI Law Society (RAILS). *J*. 2021;4(4):589-603. doi:10.3390/j4040043.
18. Ngcobo M. The ethics and law of medical AI in South Africa: balancing innovation with responsibility. *S Afr Med J*. 2025;115(5B):75-79. doi:10.7196/SAMJ.2025.v115i5b.3667.
19. Rehm HL, Page AJ, Smith L, Adams JB, Alterovitz G, Babb LJ, Barkley MP, Baudis M, Beauvais MJ, Beck T, et al. GA4GH: international policies and standards for data sharing across genomic research and healthcare. *Cell Genomics*. 2021;1(2). doi:10.1016/j.xgen.2021.100029.
20. Sarvani V, Elisha RP, Nama S, Pola LM, Rao CB. Process validation: an essential process in pharmaceutical industry. *Int J Med Chem Anal*. 2013;3(2):49-52.