WORLD JOURNAL OF PHARMACEUTICAL RESEARCH
SJIF Impact Factor 8.453
Volume 13, Issue 22, 379-399. Review Article ISSN 2277–7105

REVOLUTIONIZING CLINICAL RESEARCH WITH ARTIFICIAL INTELLIGENCE: STRATEGIES AND APPLICATIONS

Gautham J. B.* and Sudhamshu K. Tantry

Department of Pharmacy Practice, Srinivas College of Pharmacy, Valachil, Post Farangipete, Mangalore-574143, Karnataka, India.

Article Received on
26 September 2024,
Revised on 16 October 2024,
Accepted on 06 Nov. 2024
DOI: 10.20959/wjpr202422-34608

***Corresponding Author**
Gautham J. B.
Department of Pharmacy Practice, Srinivas College of Pharmacy, Valachil, Post Farangipete, Mangalore-574143, Karnataka, India.

ABSTRACT

The integration of artificial intelligence (AI) into clinical research is revolutionizing traditional methodologies, enhancing efficiency, and improving patient outcomes. This review explores the diverse applications of AI throughout the clinical research continuum, from preclinical studies to post-market surveillance. AI plays a crucial role in preclinical research by employing machine learning algorithms to identify potential drug candidates and optimize experimental designs, significantly reducing time and costs. In the clinical trial phase, AI facilitates patient recruitment through predictive analytics that identify suitable participants based on eligibility criteria, demographics, and historical data. AI-driven tools enhance patient retention strategies by enabling personalized communication and monitoring, ensuring adherence to protocols and improving trial outcomes. The integration of AI in clinical trial conduct highlights automated data collection, real-time monitoring, and adaptive trial designs that allow for dynamic modifications based on interim results. Advancements in clinical data management, enabled by AI technologies, are transforming data handling processes. Software solutions such as Oracle Clinical, SAS Clinical, and OpenClinica streamline data entry and analysis, ensuring data integrity and compliance with regulatory standards. Despite numerous benefits, challenges remain in the widespread implementation of AI in clinical research. Issues such as data privacy concerns, the need for robust validation processes, and regulatory frameworks must be addressed to facilitate the adoption of AI technologies. The potential of AI to transform clinical research is immense, offering innovative solutions to improve methodologies and patient care.

KEYWORDS: Artificial intelligence, Clinical trial, Clinical research.

INTRODUCTION

The majority of medications take ten years or longer to reach the market, cost billions of dollars, and have the potential to completely destroy a company if they fail late-stage studies after generating a lot of conjecture. Advance artificial intelligence (AI), a concept that is becoming increasingly important in addressing these problems and appears to be the key to successful drug development. In just a few minutes, IBM's artificially intelligent supercomputer, Watson, can deconstruct something that would take human researchers weeks or months to do. Watson and similar systems demonstrate progress and speed up the process by incorporating fresh results into their algorithms through machine learning.[1] AI is essential to the pharmaceutical industry at many levels, from product management to medication development. In drug screening and design, artificial intelligence (AI) technologies including machine learning (ML), deep learning, and AI-based quantitative structure-activity relationship (QSAR) technologies are essential. Data processing in drug development is facilitated by neural networks, such as recurrent neural networks (RNNs), deep neural networks (DNNs), and artificial neural networks (ANNs). SVMs help with input data regression and classification. Using AI in clinical trials helps with subject selection, lowers dropout rates through careful observation, and uses machine learning for a number of purposes. Several applications in CTs are being considered and are starting to be investigated in practice as AI is acknowledged as a route to sustainable and optimised drug development. Large and complicated volumes of categorised and uncategorised clinical, molecular, and imaging data are made available to medical research through the growth and extension of randomised studies. In order to efficiently speed up and streamline the various activities within drug research, comprehensive AI models that have been developed and trained with the right datasets are necessary to generate actionable insights from the available information, even though data availability is crucial for data-driven and personalised medicine trends.

Literature has already recognised a number of prospects, beginning with AI's role in discovery in fields where profitability may not be supported by return on investment (rare diseases, tailored medicines). Additionally, expected improvements in protocol design and patient recruiting efficiency are recommended to raise the likelihood that a trial will be successful, and AI-powered patient monitoring and analysis could have a beneficial effect on measurement and result interpretation.[2]

AI has the potential to improve traditional clinical trial design by using knowledge discovery, data mining, and NLP. These techniques can analyze diverse sources like electronic health records and medical literature to enhance patient recruitment, optimize patient-trial matching, and enable automated monitoring during trials. For instance, researchers used NLP and scalable methods to develop the Clinical Trial Knowledge Base, which standardizes eligibility criteria from over 350,000 trials listed on ClinicalTrials.gov.Even though artificial intelligence (AI) for clinical trials is making great strides, it is important to recognise the responsibility that comes with the chance to use AI to revolutionise the clinical trial cycle. According to an analysis of pertinent regulatory documents, authorities embrace artificial intelligence (AI) as a cutting-edge technology with a wide range of uses and disruptive potential, but its deployment necessitates careful planning.[3]

Utilization of ai in preclinical research

AI can help meet unmet medical requirements by improving and speeding up the discovery of new molecular targets, such as genes or proteins. Developing and training efficient and dependable algorithms that produce new stable compounds with genuine treatment potential requires access to sizable pharmacokinetics (PK) and pharmacodynamics (PD) datasets from prior preclinical and clinical research, including unsuccessful studies. One major obstacle to realising the full potential of AI in new drug discovery is the absence of published PK/PD data, which can be due to proprietary or competitive reasons.

A number of AI techniques for predicting safety are explained. Actually, there is software that uses target knowledge to anticipate medication toxicity. In vitro and animal models may eventually be replaced by effective toxicity predictions as the conventional pre-clinical method. Additionally, by offering early indication of high-risk drugs highlighted with serious safety concerns, the models can be employed as tools for risk management and prioritisation in development pipelines.

Model interpretation can be difficult, as in other areas of AI, particularly in the early stages of research when there is a lot of uncertainty. Interpretability and prediction certainty depend on an understanding of the biological systems and model properties.[2]

There are a number of solutions that the top biopharmaceutical companies have found. Pfizer is looking for immuno-oncology medicines with the aid of IBM Watson, a machine learning-based system. Nearly all of the big biopharmaceutical companies have similar internal

programs or partnerships.AI and machine learning will usher in a new era of drug discovery that is quicker, less expensive, and more efficient if proponents of these tactics are right. Although some people are dubious, the majority of specialists think these tools will become increasingly important in the future. This change presents opportunities as well as challenges for scientists, especially when the methods are combined with automation.[4]

Enhancing patient recruitment through ai

Clinical trial patient enrolment is still difficult. AI can be used to find suitable participants and match those patients with clinical trials. Natural language processing technologies, which are a collection of algorithms created to allow computers to comprehend, interpret, and produce human language, are used by clinical trial matching systems or services to learn patient RWD (real-world data) and clinical trial procedures. In order to determine a patient's eligibility, the systems are able to retrieve important information. Based on the trial eligibility requirements, a machine-learning method is used to automatically match patients to clinical trials. The system assists in prioritising and reducing the number of pertinent clinical trials to a more manageable number of trials for which the patient seems eligible. AI-based clinical trial matching systems enable accurate and efficient screening of cancer patients for clinical trials, according to a number of prior studies. AI offers a novel way to make it easier for patients to enrol in clinical trials, but it would be hard to compare the robustness of the tools against one another unless there is a clear definition and consistency on what constitutes the gold standard data for evaluating such tools.[5]

It has been emphasised in various therapeutic domains that AI systems can integrate data, including imaging, laboratory, demographic, and other -omics data, to match patients with those intricate inclusion criteria, guaranteeing their fitness for recruitment. The adoption of standardised terminology for eligibility requirements, which permits system interoperability, is a fundamental premise for the successful contribution of these AI technologies in hiring. For the tool to be used as intended, it must be able to read and comprehend the input. In order to complement information for eligibility screening, it is therefore proposed that it may be advantageous to integrate structured data with insights gleaned from natural language processing of patient reports. AI could also provide a fast and facilitated approach to internal Investigator ranking, which would speed up site initiation and therefore positively impact the recruitment process.[2] AI systems have a number of benefits that make recruiting for clinical trials easier. It has been shown that AI can automatically create prediction models that are

suited to particular issues; Convolutional Neural Networks (CNNs) are particularly good when working with larger datasets, which increases the efficacy of recruitment. The workload can be greatly reduced by using tools like CogStack, which can reduce the manual examination of patient records by 85%. AI-powered patient identification in EHRs (Electronic Health Records) reduces manual checks, expedites the hiring process, and enhances accessibility by enabling remote consent and interventions.[3]

Strategies for patient retention supported by ai

In clinical trials, about 30% of participants leave before the study is over. As a result, it is essential to offer patients ongoing assistance during their involvement. Lack of frequent communication, insufficient evaluation of patient status, and failure to update patients on trial progress—all communication-related problems—are major contributors to this dropout rate. In order to answer their enquiries, evaluate their status, and assist in managing any difficulties or side effects they may encounter, patients need continuous access to professionals. Conventional support techniques that depend on medical personnel are frequently too expensive to continue. Nearly 40% of patients stopped taking the drug which they prescribed during the first year, according to a review of 95 clinical trials.[6]

While high nonadherence rates provide significant obstacles for sponsors and researchers, underpowered trials may arise from a lack of technical infrastructure to handle trial complexity without trustworthy adherence control and patient monitoring. There was no research on using AI to address the retention issue, despite its severity, according to our review. Using machine learning techniques, a conceptual framework for remote health monitoring in a clinical trial was created. It included multiple classifications to decide whether or not a participant should be permitted to continue in the experiment, but it did not validate the framework's impact. Numerous AI uses in clinical trial settings were covered in a recent review study. These applications include the use of AI technology to analyse video recordings of patients taking medications in order to ensure proper dosages and the use of historical medical data to predict potential patient dropouts so that medical professionals can intervene promptly. With its unique product that processes both structured and unstructured patient data before sending customised messages and alerts intended to maintain participants' involvement in trials while also identifying individuals at risk of dropping out for prompt intervention, Brite Health stands out as a company actively addressing clinical trial dropout

rates at a time when there aren't many startups concentrating specifically on preventing clinical trial dropouts and engaging participants.

Furthermore, digital techniques known as directly observed therapy (DOT), in which an AI system tracks patients' medicine intake, are used by companies like emocha Mobile Health and AiCure to improve medication adherence. Even modest improvements have required remarkably complicated tactics, while simple actions frequently have little effect. Additionally, all intervention effects tend to wane with time in the absence of ongoing reinforcement strategies, emphasising the necessity of comprehensive retention strategies that include patient-centered care, education, counselling, dosage monitoring, timely packaging, and reminders. There are many chances for innovation in this field because patient dropout is probably going to be a major issue for some time to come.[3]

Integrating ai in the conduct of clinical trials

In addition to improving productivity and clinical development results, artificial intelligence may be able to reduce the length of clinical trial cycles. Real-world data (RWD) is the term used to describe the growing amounts of scientific and research data that biopharma companies can now access from a variety of sources. However, they have often lacked the knowledge and resources necessary to make effective use of this data. Researchers may discover pertinent patients and key investigators, improve their understanding of diseases, and create ground-breaking clinical study designs by using predictive AI models and advanced analytics to uncover RWD.AI algorithms may be used to clean, aggregate, code, preserve, and retain clinical trial data when paired with an effective digital infrastructure. Furthermore, improved electronic data capture (EDC) may reduce the impact of human error in data collection while also allowing for smooth system integration.[4]

The integration of automated data collection tools and AI-driven digital biomarkers allows investigational sites near real-time access to participant data through wearable devices and sensors, providing visual insights into a participant’s condition. This enhances safety monitoring, particularly for those with life-threatening or severe conditions, by offering quicker access to actionable data. For example, a video capture device with an AI algorithm can reliably confirm medication adherence, offering a more practical alternative to direct observation by site staff, which is often impractical.[2]

Unlearn has begun collaborating with academic institutions, biotech firms, and sponsors of pharmaceutical corporations to optimise TwinRCTsTM, a clinical trial software. To improve the success rate of trials with smaller patient groups, TwinRCTsTM combines artificial intelligence (AI), digital twins—a virtual model designed to accurately replicate a physical entity—and advanced statistical techniques. The AI model will use past control data to create a digital twin for every patient, in contrast to conventional clinical trials. In external cohorts, the digital win can forecast the course of the disease. By comparing the patient with his digital twin, it is possible to properly forecast how the treatment would affect the primary and secondary endpoint scans. The European Medicines Agency's draft opinion, which proposes that this technique can be used in primary analyses of phase II and III clinical trials because it does not create bias, demonstrates the possible policy consequences of this. Because patients in TwinRCTsTM are more likely to get potentially helpful treatment rather than a placebo, the smaller control group is more appealing to them than in traditional trials. Because they may spend less time enrolling patients and meeting enrolment goals, it is also more attractive to other stakeholders.[5]

Advancements in clinical trial design through ai

Ineffective patient cohort selection and recruitment strategies, as well as the incapacity to adequately follow patients throughout trials, are two of the primary reasons for high trial failure rates1.In recent years, our approach to clinical trials has started to evolve thanks to AI-enabled technology and real-world data (RWD)—scientific data gathered from diverse sources within the healthcare sector. This advancement enables us to reshape key aspects of clinical trial design.[5]

When AI techniques are incorporated into the design, better results are seen. For instance, protocol enrichment and biomarker verification enhance the cohort composition's appropriateness. However, in order to reduce inaccuracy in AI output, it is necessary to work together to develop common procedures for gathering, archiving, and organising huge datasets.[7] Furthermore, in a virtual control arm, well-designed AI algorithms with access to sufficient high-quality data could be utilised to forecast the progression of a disease. This might make it possible to substitute a totally virtual arm with only synthetic data for a placebo arm.[8] A number of advantages are anticipated, such as lower costs, less strain on the site and patients, and maybe quicker CT scans.

However, it takes a lot of effort and money to validate virtual control arms based on training datasets created alongside current CTs. Additionally, synthetic control arms could allay ethical worries about placebo control groups and persuade volunteers who might not otherwise be prepared to take the chance of being randomly assigned to a placebo.[7] From planning to implementation, data-driven AI solutions have enormous potential to enhance the clinical trial design process. By speeding up patient-to-trial matching and enrolment and dynamically monitoring patients throughout trials, it raises the likelihood of success and can enhance adherence control and produce more accurate and meaningful endpoint evaluations.

Nonparametric Because of its adaptability, Bayesian learning has become a potent tool in contemporary machine learning frameworks, offering a Bayesian framework for nonparametric model selection. Currently, one of the most popular Bayesian nonparametric models is the Dirichlet process, especially in Dirichlet process mixture models (Sometimes called infinite mixture models). Without pre-specifying the number of components in a mixture model, Dirichlet process mixtures offer a nonparametric method for determining model densities and locating latent clusters within the observed variables. These Bayesian nonparametric models have numerous uses in clinical trial design. Nonparametric Bayesian learning, for instance, can provide effective and efficient dose selection in oncology dose-finding clinical trials.

In oncology, patients with various cancer kinds are frequently enrolled in human trials, leading to heterogeneity. These problems may be particularly noticeable in cell treatments and immuno-oncology. Inaccurate dose selection and ineffective target population identification may result from designs that disregard the variability of safety or efficacy profiles across different tumour types. These Bayesian logistic regression model (BLRM)-based designs enhance the effectiveness of the dose search and the precision of estimating the ideal dose level by enabling data-driven borrowing of information across several populations while taking heterogeneity into consideration. These trials have a greater chance of producing efficacy and safety data more quickly through adaptive decision-making when nonparametric Bayesian learning is used.[9]

Ai in clinical data management

Clinical Data Management (CDM) is defined as the systematic collection, integration, and validation of clinical trial data. When regulatory agencies evaluate new drugs, they rely heavily on the accuracy and integrity of this data. Confidence in the clinical data is typically

tied to the quality practices and standards maintained throughout the trial. Therefore, the primary goal of CDM is to ensure high-quality data by minimizing errors and missing entries, while maximizing data collection for comprehensive analysis.

Various sectors, including medical research, biotechnology, and pharmaceuticals, benefit from clinical trial software that supports all phases of a clinical trial—from participant recruitment to study submission and archiving. These software solutions are versatile and can be used for both small Phase I trials and large-scale studies with thousands of participants. Known as Clinical Data Management Systems (CDMS), these tools include commercial options such as ORACLE CLINICAL, MACRO, RAVE, CLINTRIAL, eClinicalSuite, and EZentry. Despite their effectiveness, commercial CDMS tools can be costly and require sophisticated IT infrastructure to function efficiently.

In addition to commercial tools, there are open-source CDMS options like OpenClinica, TrialIDB, openCDMS, and PhOsCo. These tools offer several advantages, such as accelerating study timelines, controlling costs, providing accurate trial predictions, supporting robust program modeling, and assisting with screening, scheduling, recruiting, and even adjusting for foreign exchange rates.[10]

1. EZ entry Systems

A software application called EZ entry was created using changes made to the EpiData software system. The primary features of the EpiData software system were 1) the use of straightforward syntax for e-CRF setup 2) Data entry combined with checkout theory 3) Verification of the data following double entry 4) Data export in multiple formats. EpiData's primary flaw was that it was a low-security system. EZ entry includes query management, revision tracking, data entry, and data import and export. EZ-entry serves two primary purposes: quality control and security protection. User authentication and revision tracking are the two components that make up the security of the EZ-entry system. Only authorised users are granted access to the database under the user authentication mechanism. When an authorised user makes a revision to the original database or adds new data, the system immediately logs the action in the revision tracking. Field value check, data entry alignment, and query form are some of the aspects that affect the quality of the data in the EZ-entry system.[10]

2. Oracle clinical software

Experts have been using Oracle software for 30 years since it offers information with reliable, integrated, and secure technologies. The Oracle software offers a number of benefits, such as improved teamwork, quicker installation, successful marketing, and a higher return on investment. More than 200 companies that deal with pharmaceutical and biotechnology products have used Oracle Clinical to perform more than 10,000 clinical trials.

Key features of the software include

- Top-notch remote data collection and clinical data management services.
- A smooth transition from paper to electronic trial data collection.
- A CRF with annotations.
- More audit trials to guarantee industry compliance;
- Better randomisation;
- Study data that is easy to understand

3. TCS CLIN-E2E software

All four stages of clinical trials are covered by this software. The software's safe electronic environment records the electronic data that connects the laboratories and sites to the sponsors.

Key features of the software include

- It is powered by the aforementioned protocol.
- Able to manage CRFs in both paper and computer formats.
- Produces reusable research templates and CRFs, cutting down on study time. The software performs a number of tasks, including:
- Data correction and modification.
- Complies more closely with GCP part 11 of 21 CFR.

4. SAS Clinical software

Key features of the software include

- Easy-to-use, integrated systems that enable data management and access from a variety of sources, such as CDMS, EDC, etc.
- It reduces the need for manual intervention by mechanising and detecting various operations.

- The appropriate use of data ensures that standards are applied correctly. SAS helps to manage and automate data integration and achieve data quality that complies with industry standards, freeing up valuable resources to work on more complex topics and saving time.[10]

5. Cognos Business Intelligence software

24 of the top 30 pharmaceutical companies in the world utilise Cognos, one of the biggest suppliers of business presentation software. Today, Cognos has over 23,000 clients, including a wide variety of health care and life sciences organisations. In the past, point-solution systems with limited decision assistance, such as EDC and CTMS, were used. By precisely assessing the data quality, employee and partner performance, and many other factors, Cognos has simplified data management.[11] This program offers service-oriented architecture (SOA) along with a range of business intelligence features.

6. Akaza's OpenClinica software

It provides an open informatics solution which keeps a record of the needs of the organizations that are involved in research.

Key features of OpenClinica include

- Clinical research that is properly organised according to the site and methodology.
- Provides transparent and safe resource sharing.
- Creation of online CRFs.
- Appropriate data management, including regular patient visits.
- Data import and export tools to transfer clinical datasheets between Excel spreadsheets.
- Interfaces for retrieving and querying data[10]

Clinical Data Management (CDM) has evolved from paper-based to electronic systems due to the growing need for expedited drug development, which has greatly improved the speed and quality of data gathering. While CDM specialists concentrate on upholding standards, regulatory bodies make sure that reliable quality systems generate high-quality data. Assuring regulatory compliance and efficiently designing and implementing data management systems are the two biggest problems in CDM. Despite these obstacles, technological developments have made it possible for CDM to develop into a standard-based organisation that strikes a balance between business needs and current system limitations.[10]

The role of machine learning in clinical research

By being exposed to data, a mathematical model known as machine learning (ML) can become more proficient at a task. Preclinical research and observational studies lead to traditional trials and trials with pragmatic features, which in turn lead to clinical registries and additional implementation work. Clinical research is a broad subject. Clinical research as it is now carried out is complicated, time-consuming, costly, and susceptible to biases and unforeseen errors that can occasionally jeopardise its successful application, implementation, and acceptance, despite the fact that it is essential to enhancing healthcare and results. Clinical trial success, generalisability, patient-centeredness, and efficiency could all be enhanced by machine learning. ML can benefit clinical trials in a variety of ways, from preclinical drug discovery to pre-trial planning to study execution.

1. Machine learning in preclinical drug Discovery and Development

Promising candidate compounds and targets are discovered, the exploratory strategy to obtain regulatory approval is developed, and extensive preclinical research and preparation are necessary for successful clinical trials. Errors in this stage could cause clinical studies to fail in the end or postpone the discovery of potential medications. Researchers can reduce the inefficiencies of the preclinical process by utilising both past and current research with the aid of machine learning.[12]

By synthesising vast volumes of current research, elucidating pharmacological processes, and predictively modelling protein structures and future drug target interactions, machine learning (ML) can expedite and improve the success of therapeutic target identification and candidate molecule development. By optimising molecules within the limitations of a target biological system, for example, using a gated graph neural network, machine learning (ML) can also help generate potential candidate compounds.[13] Applying a Bayesian machine learning technique to an anti-cancer compound demonstrated how ML can synthesise and analyse vast volumes of data to better understand the medication's mechanism in circumstances where a drug candidate performs differently in vivo than anticipated.[14]

2. Optimizing clinical study protocols with machine learning

By using simulation techniques to vast amounts of data from previous trials, machine learning (ML) can help maximise trial success and efficiency during the planning phase as medicinal compounds get closer to human trials. This will make it easier to build trial protocols. As demonstrated by reinforcement learning techniques to Alzheimer's disease and

non-small cell lung cancer, for example, study simulation may optimise the selection of treatment regimens for testing.[15,16] AI employs natural language processing to find potential problems and obstacles to a successful trial outcome and enables investigators to upload protocols. Preclinical research and clinical trial preparation can be made more effective and productive with the application of machine learning. However, most peer-reviewed reports of ML use in this capacity focus on preclinical research and development rather than clinical trial planning.

3. Machine learning in clinical trial participant management

By evaluating clinical trial data sets "to predict which patients will respond to a particular therapy in development, thereby improving inclusion/exclusion criteria and ensuring primary study outcomes are achieved," Bullfrog AI, a start-up, has attempted to leverage the potential of targeted patient population selection.[17] The time-consuming procedure of manual selection is lessened by natural language processing (NLP), which provides an effective way to find patients who meet particular characteristics. Compared to conventional machine learning techniques, a cross-modal inference learning model improves patient-trial matching utilising EHR data by concurrently encoding enrolment criteria and patient records. Similar services are offered by commercial systems like Mendel.AI and DeepAI, albeit there is a lack of peer-reviewed validation. NLP can reduce cohort bias by avoiding the need for structured data, but it may still introduce biases from the EHR documentation itself, potentially affecting algorithm performance.[12]

4. Enhancing participant Retention, Monitoring and Protocol adherence

Two overarching machine learning (ML) strategies are emerging to improve participant retention and adherence in clinical trials. The first strategy involves using ML to analyze large amounts of data to identify participants at risk of non-compliance, enabling proactive interventions. The second aims to reduce participant burden, improving their overall trial experience by streamlining data collection through methods like passive monitoring or repurposing clinical and real-world data. For example, AiCure uses facial recognition technology to track medication adherence, though concerns remain about privacy and validation across diverse patient groups. ML can also extract valuable insights from clinical documentation, wearable devices, and even social media, reducing the workload for participants. While these innovations hold significant potential, challenges remain around

privacy, bias, and the need for rigorous validation, with further studies required to assess their impact on clinical trial quality, diversity, and participant engagement.[12]

5. Data Collection and Management via Machine Learning

The methods needed for data collection, management, and analysis may alter when machine learning is used in clinical trials. Automating data collecting into case report forms is an enticing use of machine learning (ML), more especially natural language processing (NLP), to research data management. This reduces the time, cost, and error-proneness of human data extraction, whether in prospective trials or retrospective reviews. Risk-based monitoring techniques for clinical trial surveillance can be powered by machine learning (ML), which makes it possible to avoid and/or detect fraud, site failure, and missing or inconsistent data that could postpone database locking and further analysis. Using auto-encoders to separate plausible from implausible data might help detect suspicious data trends in clinical trials or inaccurate data in observational studies. Clinical trial data processing can be streamlined by machine learning (ML), especially in semiautomated endpoint identification and adjudication, which lowers the complexity, expense, and time involved in doctors' manual event adjudication. Although these techniques lack peer-reviewed validation, businesses such as IQVIA have used optical character recognition and natural language processing (NLP) to demonstrate ML's capacity to categorise and categorise events, such as adverse drug responses. One major issue is that endpoint definitions vary from trial to trial, necessitating model retraining for every study. Efforts are being made to standardise results in cardiovascular research, however there are currently few efforts to aggregate trial data for model training, and comparable programs are not present in other domains. Consensus on event definitions and data sharing across several trials for efficient model training will be essential for future developments.[12]

6. Data analysis in clinical research

ML is ideally suited for hypothesis development, risk modelling, and counterfactual simulation, all of which can be facilitated by data gathered from clinical trials, registries, and clinical practices. In real-world data, for example, unsupervised learning can find phenotypic clusters that might be investigated further in clinical trials.[18,19] Additionally, by more effectively detecting treatment heterogeneity while still offering some (albeit insufficient) protection against false-positive findings, machine learning (ML) has the potential to enhance the widely used practice of secondary trial analyses and reveal more exciting directions for

further research.[20,21] Furthermore, risk predictions in retrospective datasets that can later be prospectively confirmed are efficiently produced using machine learning. In conclusion, there are numerous efficient machine learning algorithms for managing, processing, and analysing clinical trial data, but there are less methods for enhancing the calibre of data as it is created and gathered. Since ML techniques rely on the availability and quality of data, conducting high-quality trials is still crucial to enabling higher level ML processing.[12]

Artificial intelligence application in cancer research

AI has been used to cancer research for almost 20 years, and new developments have showed a lot of promise, performing at an expert level in a number of domains. Clinical research on cancer has seen notable advancements thanks to AI, especially deep learning (DL), and the growing availability of rich biomedical data. These days, AI-based techniques are widely used in many different domains, such as drug discovery, genetic analysis, medical record mining, cancer imaging, and the use of biomedical literature.

AI aids radiologists in diagnosing illnesses, while image preprocessing methods such as lesion segmentation increase precision. AI is essential to drug research because it helps create candidate compounds based on 3D protein structures, improving toxicity prediction and bioavailability. Finding genetic biomarkers predictive of treatment response is another benefit of molecular profiling from tissue samples. Platforms have developed through collaborations between pharmaceutical and IT companies aim to improve biomarker discovery and drug sensitivity predictions.

Large public databases have facilitated the advancement of cancer drug research and improved the accuracy of anticancer treatment response forecasts. While DeepChem, DeepTox, AlphaFold, and DeepNeural NetQSAR are used for drug discovery and toxicity predictions, lowering costs in the process, CNNs and other AI models have been developed to predict drug efficiency at the single-cell level. Not withstanding AI's advancements, there are still obstacles to overcome, especially in acquiring sizable, thoroughly annotated cancer datasets. Cloud computing, sophisticated prediction models, and data sharing agreements can all aid in overcoming these obstacles. The quality of data, which still needs to be standardised, is critical to the effectiveness of AI in cancer research. AI is anticipated to be a key component of cancer treatment in the future, enhancing diagnosis, treatment outcomes, survival rates, and minimising adverse effects.[22]

Barriers to the implementation of ai in clinical research

AI also has several limitations that could impede its effectiveness in clinical trial recruitment, such as the "black box" issue, which raises concerns about transparency. While AI can improve resource efficiency in recruitment, it may unintentionally introduce biases into trial cohorts. For instance, AI tools might limit the applicability of models like A4 to younger participants. Additionally, issues related to the quantity and quality of training data, as well as integrating AI frameworks into existing systems, pose significant challenges, especially since these tools have not been widely tested in real-world settings. Consumer-oriented search engines for trial registries or patient-trial matching systems, along with the generalizability of these methods to other common diseases, remain underexplored. Studies have also highlighted problems related to fairness, discrimination, and selection bias, as machine learning models may inadvertently reflect biases present in their training data, resulting in unfair outcomes. RWD's uneven quality, data sharing difficulties, and the absence of standards for the ethical, repeatable, scalable, and explainable incorporation of AI into clinical trials. The foundation of AI models is high-quality data, which necessitates the creation of standardised biomedical databases that include data from wearables, health applications, clinical records, medical imaging, omics, and social media. Data privacy regulations and intense competition between organisations make data sharing difficult. A more cooperative attitude within and across institutions, along with the use of privacy-protecting technologies like swarm learning and data encryption, can change this situation.[5]

Implementing machine learning (ML) processing of device data is hampered by the need to better define the overlap between patient-centric digital biomarkers and previously established clinical endpoints, as well as by the need to comprehend participant privacy attitudes regarding the sharing and use of device data. Novel biomarkers will also need FDA approval. If not handled effectively and openly, operational obstacles to ML integration in clinical research can exacerbate and strengthen philosophical worries. Racial bias in model application, for example, can result from limited training data and poor model calibration, as has been observed in ML formelanoma identification.[12] Even though AI frequently performs well in medical research, its use in real-world scenarios is limited due to the model's somewhat opaque nature. The machine could explain how it learnt and why it arrived at this conclusion. This is known as the "blackbox" problem because it is challenging to show which aspects of the input data contribute to the output. For instance, AI can predict the best course of treatment for a patient without providing the reasoning behind the prediction.

InterpretableDL is a trendinal lowering this limitation.[22] In order to ensure that AI is applied appropriately and practically to satisfy clinical demands, human expertise and involvement are crucial, and the absence of this expertise may be detrimental to the practical application of AI. It will take a multidisciplinary approach, creative data annotation tools, and the creation of more exacting AI models and procedures to address these issues and offer constructive solutions.

Current landscape of clinical research utilizing ai techniques

The topic of artificial intelligence (AI) has attracted significant attention recently and is predicted to transform the health-care landscape in the near future. It includes several analytical techniques like computer or machine vision, natural language processing, and machine learning (ML), including deep learning. According to the surveys, the number of AI-related research projects has been steadily rising over time. This study's most frequent topic of interest was oncology, with breast cancer serving as its primary emphasis. "CURATE. AI" is the most recent artificial intelligence platform to be developed. It has been effectively applied to patients with metastatic castration-resistant prostate cancer and is being extended to other cancer types.

Similarly, researchers at Google AI Healthcare created Lymph Node Assistant (LYNA), a learning system designed to analyze histology slides containing tissue samples from lymph node biopsies. This could help diagnose metastatic breast cancer tumours and spot problematic areas that are invisible to the naked eye. In addition to demonstrating a 99% accuracy rate in identifying a sample as either malignant or noncancerous, LYNA also cut the typical slide viewing time in half. Similarly, researchers at Google AI Healthcare developed LYmph Node Assistant (LYNA), a learning algorithm that examined histology slides stained with tissue samples from lymph node biopsies. This could help diagnose metastatic breast cancer tumours and spot problematic areas that are invisible to the naked eye.

The average slide review time was cut in half by LYNA, which also demonstrated a 99% accuracy rate in identifying a sample as either malignant or noncancerous. Since 2017, the number of studies utilising AI has skyrocketed, with over one-third of all studies being registered in 2019 alone. Artificial Intelligence is already being used in medicine, from drug dosage algorithms to online check-ins at medical centres, appointment and follow-up scheduling, medical record digitisation, and child and pregnant female immunisation dates.

As a result, the expansion of this field of study will soon open up new avenues for medicine.[23]

The future of ai in clinical Research and Healthcare

In order to provide effective, mobile, real-time, continuous, and customised patient surveillance systems that can efficiently monitor patients during the trial period and minimise site visits, artificial intelligence approaches may eventually be integrated with smart devices, such as wearable sensor devices. For example, a Stanford University research team has created a flexible electronic strain sensor for in vivo monitoring of the dynamic change in tumour volume, marking a significant advancement in wearable technology and offering a promising method to reflect therapeutic success.[24] Furthermore, the systems have the ability to automatically and continuously gather, process, and manage patient data in order to identify patients who are likely to suffer from severe adverse events, detect or create clinical endpoints free from human bias, and estimate a patient's risk of dropping out.[5] AI has been shown to mimic human knowledge in computer innovations that could benefit patients and specialists in the following ways:

- By providing a research facility for the analysis, depiction, and categorisation of restorative data.
- Developing new tools to support research and decision-making; and, finally, adding activities related to programming, psychology, and medicine.
- By providing a substantively rich framework for a logical restorative group in the future.

In this context, the broader integration of advanced AI tools into routine medical applications could improve treatment effectiveness and reduce costs by minimizing the risks of misdiagnosis, enabling more targeted preoperative planning, and lowering the likelihood of complications during surgery.[1] The integration of AI into the pharmaceutical industry promises to revolutionize drug discovery, development, and delivery. AI-powered virtual screening will significantly accelerate the identification of therapeutic candidates by efficiently analyzing large chemical libraries. In personalized medicine, AI can categorize patients and predict therapy responses using genomic, proteomic, and clinical data, allowing for customized treatments. Furthermore, AI-driven deep learning models hold the potential to create innovative compounds with enhanced target-binding properties, improving drug efficacy while minimizing side effects. Additionally, AI algorithms will optimize patient-

specific formulations, enhance safety assessments, and predict adverse effects, leading to better treatment outcomes.[25]

CONCLUSION

The integration of AI into clinical trials is an emerging and promising area, with many experts suggesting that AI could be the key to overcoming the current challenges in drug development and ushering in a new era of sustainable medical research. The growing efforts to evaluate AI's potential demonstrate the industry's eagerness to unlock its full capabilities for more successful and cost-effective trials. AI technologies have made significant strides, and numerous companies are leveraging these advancements to address business challenges, while also breaking new ground in the process. It is crucial to maintain transparent discussions about the benefits and limitations of machine learning (ML) in clinical research, as well as to share best practices across academia, government, and public forums. This openness ensures that the application of ML in clinical research is done in a fair, ethical, and transparent way that is widely accepted. Currently, AI's use in clinical trials is still in the early stages, with few tools ready for widespread implementation. Its effectiveness needs further validation, requiring more robust, high-quality research evidence and clearer reporting. Therefore, sponsors, investigators, and regulators aiming to standardize AI's use in clinical trials must collaboratively build the necessary infrastructure and expertise to ensure patient safety and protection remain the top priority.

REFERENCES

1. Nayak VS, Khan MS, Shukla BK, Chaturvedi PR. Artificial intelligence in clinical research. Int J Clin Trials, 2016; 3(4): 187-93.
2. Askin S, Burkhalter D, Calado G, El Dakrouni S. Artificial intelligence applied to clinical trials: opportunities and challenges. Health and technology, 2023; 13(2): 203-13.
3. Lu X, Yang C, Liang L, Hu G, Zhong Z, Jiang Z. Artificial intelligence for optimizing recruitment and retention in clinical trials: a scoping review. Journal of the American Medical Informatics Association, 2024; 31(11): 2749-59.
4. Shaheen MY. Applications of Artificial Intelligence (AI) in healthcare: A review. ScienceOpen Preprints, 2021; 25.
5. Zhang B, Zhang L, Chen Q, Jin Z, Liu S, Zhang S. Harnessing artificial intelligence to improve clinical trial design. Communications Medicine, 2023; 21, 3(1): 191.

6. Hutson M. How AI is being used to accelerate clinical trials. Nature, 2024; 1, 627(8003): S2-5.
7. Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. Trends in pharmacological sciences, 2019; 1, 40(8): 577-91.
8. Lee CS, Lee AY. How artificial intelligence can transform randomized controlled trials. Translational vision science & technology, 2020; 28, 9(2): 9.
9. Kolluri S, Lin J, Liu R, Zhang Y, Zhang W. Machine learning and artificial intelligence in pharmaceutical research and development: a review. The AAPS journal, 2022; 24: 1-0.
10. Kaur S, Singh I. Artificial intelligence based clinical data management systems: A review. Informatics in Medicine Unlocked, 2017; 1, 9: 219-29.
11. Almenoff JS, Pattishall EN, Gibbs TG, DuMouchel W, Evans SJ, Yuen N. Novel statistical tools for monitoring the safety of marketed drugs. Clinical Pharmacology & Therapeutics, 2007; 82(2): 157-66.
12. Weissler EH, Naumann T, Andersson T, Ranganath R, Elemento O, Luo Y, Freitag DF, Benoit J, Hughes MC, Khan F, Slater P. The role of machine learning in clinical research: transforming the future of evidence generation. Trials, 2021; 22: 1-5.
13. Liu Q, Allamanis M, Brockschmidt M, Gaunt A. Constrained graph variational autoencoders for molecule design. Advances in neural information processing systems, 2018; 31.
14. Madhukar NS, Khade PK, Huang L, Gayvert K, Galletti G, Stogniew M, Allen JE, Giannakakou P, Elemento O. A Bayesian machine learning approach for drug target identification using diverse data types. Nature communications, 2019; 19, 10(1): 5221.
15. Romero K, Ito K, Rogers JA, Polhamus D, Qiu R, Stephenson D, Mohs R, Lalonde R, Sinha V, Wang Y, Brown D. The future is now: model-based clinical trial design for Alzheimer's disease. Clinical Pharmacology & Therapeutics, 2015; 97(3): 210-4.
16. Zhao Y, Zeng D, Socinski MA, Kosorok MR. Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. Biometrics, 2011; 67(4): 1422-33.
17. Weissler EH, Naumann T, Andersson T, Ranganath R, Elemento O, Luo Y, Freitag DF, Benoit J, Hughes MC, Khan F, Slater P. The role of machine learning in clinical research: transforming the future of evidence generation. Trials, 2021; 22: 1-5.
18. Ngufor C, Warner MA, Murphree DH, Liu H, Carter R, Storlie CB, Kor DJ. Identification of clinically meaningful plasma transfusion subgroups using unsupervised random forest clustering. InAMIA Annual Symposium Proceedings, 2017; 2017: 1332. American Medical Informatics Association.

19. Tomic A, Tomic I, Rosenberg-Hasson Y, Dekker CL, Maecker HT, Davis MM. SIMON, an automated machine learning system, reveals immune signatures of influenza vaccine responses. The Journal of Immunology, 2019; 1, 203(3): 749-59.
20. Watson JA, Holmes CC. Machine learning analysis plans for randomised controlled trials: detecting treatment effect heterogeneity with strict control of type I error. Trials, 2020; 21: 1-0.
21. Rigdon J, Baiocchi M, Basu S. Preventing false discovery of heterogeneous treatment effect subgroups in randomized trials. Trials, 2018; 19: 1-5.
22. Shao D, Dai Y, Li N, Cao X, Zhao W, Cheng L, Rong Z, Huang L, Wang Y, Zhao J. Artificial intelligence in clinical research of cancers. Briefings in Bioinformatics, 2022; 23(1): 523.
23. Karekar SR, Vazifdar AK. Current status of clinical research using artificial intelligence techniques: A registry-based audit. Perspectives in clinical research, 2021; 1, 12(1): 48-52.
24. Abramson A, Chan CT, Khan Y, Mermin-Bunnell A, Matsuhisa N, Fong R, Shad R, Hiesinger W, Mallick P, Gambhir SS, Bao Z. A flexible electronic strain sensor for the real-time monitoring of tumor regression. Science Advances, 2022; 16, 8(37): 6550.
25. Bhabad S, Lamkhade D, Koyate S, Karanjkhele K, Kale V, Doke R. Transformative trends: A comprehensive review on role of artificial intelligence in healthcare and pharmaceutical research.