Generative AI in Clinical Trials: Revolutionizing Design, Analysis and Prediction in Medical Research

Ángel Canal-Alonso1, Noelia Egido1, Pedro Jiménez1, Javier Prieto[1], Juan Manuel Corchado[1]

[1]Department of Bioinformatics and Computational Biology, AIR Institute, Carbajosa de la Sagrada, Spain

Email: acanal@air-institute.com

Abstract

In the context of clinical trials, the implementation and adaptation of emerging technologies is essential to accelerate and improve medical research. This article explores the impact and applications of generative artificial intelligence (AI) at various stages of clinical trials. Its potential benefits are highlighted, such as optimizing patient selection, innovation in protocol design, accurate prediction of outcomes, and deeper analysis of complex data. Furthermore, the importance of combining these advanced tools with human knowledge and experience is highlighted, and the associated ethical and privacy considerations are mentioned. With generative AI,

Keywords: Generative Artificial Intelligence, Drug Design, Clinical trials

1. Introduction

A clinical trial is a carefully planned investigation that seeks to discover, interpret and compare the effects—whether beneficial, neutral or harmful—of medical treatments in humans. These trials are essential for the development of new medical interventions, as they allow the evaluation of the effectiveness and safety of medications, devices, therapeutic procedures and prevention regimens.

Clinical trials are divided into several phases, each with a specific objective and a particular methodology:

- Phase I: In this initial phase, the safety of a new intervention is studied in a small group of people (usually less than 100). The main goal is to determine the optimal dosage, how it is administered and how the body processes it, as well as to identify potential side effects.

- Phase II: Once safety has been established in Phase I, Phase II evaluates the effectiveness of the treatment in a larger group of people (hundreds). Safety is also continuing to be evaluated and the appropriate dosage is being sought.

- Phase III: This phase involves an even larger number of participants (from hundreds to thousands) and takes place in multiple locations. In Phase III, the treatment is compared to existing standard treatments or a placebo to determine its effectiveness and monitor side effects. If a drug or intervention shows positive results in this phase, it can be requested for approval from regulatory authorities.

- Phase IV: After a treatment has been approved and is on the market, additional studies may be conducted in Phase IV to evaluate aspects such as long-term effects, benefits in specific subpopulations, and how the treatment compares to others already approved.

With the advancement of information technologies and the emergence of artificial intelligence (AI) in multiple fields, clinical research has not remained on the sidelines. In particular, generative AI has shown to be a promising tool with the potential to revolutionize the way we design, conduct and analyze clinical trials.

Pharmaceutical companies face multiple challenges when running clinical trials, each of which can influence the efficiency, cost, and success of developing a new treatment. Below are some of the main challenges:

1. Participant recruitment and retention: Finding and keeping the right people for a trial can be extremely challenging. Pharmaceutical companies often have difficulty recruiting enough participants who meet specific criteria. Additionally, keeping these participants engaged for the entire duration of the trial can be equally challenging.

2. High costs: Drug development is an expensive process. Clinical trials in particular, with their need for monitoring, adequate facilities and specialized personnel, can require significant investments. Reducing costs without compromising quality is a constant concern.

3. Logistical complexity: Coordinating multiple aspects of a trial, from administering treatments to patient monitoring and data collection, can be a monumental task, especially for multicenter trials that take place in multiple locations or countries.

4. Rules and regulations: Companies must ensure that their trials comply with a multitude of regulations and ethical guidelines. Regulations can vary between countries and staying up to date and compliant is essential, but it can also be complicated and time-consuming.

5. Essay Design: Designing an essay that provides meaningful and useful results is an art in itself. Determining factors such as sample size, trial duration, and appropriate assessment points can be complex.

6. Data Interpretation: With the large amount of data collected in clinical trials, analyzing and interpreting this data effectively and accurately is essential. However, this can be challenging, especially when the results are ambiguous or contradictory.

7. Ethical challenges: Ensuring that trials are carried out in a way that protects and respects the rights and well-being of participants is essential. This includes ensuring appropriate informed consent, maintaining confidentiality and ensuring risks are appropriately communicated and managed.

8. Time: Time is of the essence in drug development, especially when dealing with life-threatening or debilitating diseases. However, conducting meticulous and comprehensive clinical trials takes time, which can delay promising treatments from reaching the market.

These challenges have led the industry to seek innovative solutions and, in this context, emerging technologies such as AI have great potential to offer solutions that address these and other challenges in the execution of clinical trials.

Artificial intelligence

Artificial intelligence (AI) is a branch of computing that focuses on the creation of systems capable of performing tasks that, until now, required human intelligence to be executed. These tasks include, but are not limited to, pattern recognition, logical reasoning, decision making, learning from experience, and adaptation to new stimuli or situations.

There are several approaches and techniques in AI, including:

- Machine Learning: It is a subdiscipline of AI that allows machines to learn from data. Instead of being explicitly programmed to perform a task, machines are trained using large data sets and algorithms that allow them to improve their performance over time.

- Neural networks: Inspired by the functioning of the human brain, these mathematical structures simulate the interaction of neurons and are essential for complex tasks such as image recognition and natural language processing.

- Natural language processing (NLP): It is the field of AI that focuses on the interaction between computers and human language. Its goal is to allow machines to understand, interpret and generate language in a similar way to humans.

- Computer vision: Focuses on allowing machines to interpret and make decisions based on images or videos, similar to human visual ability.

- Robotics: Although a discipline in itself, AI plays a crucial role in creating autonomous robots that can interact with their environment and make decisions based on the information they collect.

- Expert systems: These are programs that emulate the decision-making process of a human expert in a specific domain. They use knowledge bases and rules to reach conclusions or solutions.

AI has become an essential tool in many industries, from medicine to automotive to finance, due to its ability to process large amounts of data at speeds unimaginable to humans and find patterns or solutions that may not be obvious to humans. human analysts. In the context of clinical trials, AI has the potential to address many of the challenges mentioned above, optimizing processes, improving accuracy, and accelerating the development of drugs and treatments.

Generative AI is a subset of artificial intelligence that focuses on the creation, generation, or synthesis of content, data, or information that did not previously exist. While many traditional AI applications are designed to analyze and classify information (a discriminative approach), generative systems are designed to produce new examples that are consistent with an existing data set.

The main features and applications of generative AI include:

1. Generative Adversarial Networks (GANs): These are structures composed of two neural networks: a generator and a discriminator. The generating network creates new data, while the discriminator tries to distinguish between real and generated data. Both networks are trained together, with the generator improving its ability to create data that the discriminator cannot distinguish from real data.

2. Content synthesis: Generative models can be used to create all types of content, from images and videos to music and text. For example, realistic images of non-existent people or music in a particular style can be generated without human intervention.

3. Data augmentation and improvement: In fields where data is scarce, generative AI can be used to create more examples and improve the quality of existing data sets.

4. Simulation: Generative systems can simulate scenarios or data based on previous examples, which is useful in areas such as research, product design or training.

5. Design and creation: From product design to architecture, generative AI can be used to automatically explore a wide space of possibilities and propose innovative solutions based on defined parameters and constraints.

6. Personalization: In applications such as advertising or entertainment, generative systems can be used to adapt or create specific content for a particular user or audience.

In the context of clinical trials, generative AI has significant potential. For example, it could be used to simulate patient data, allowing researchers to explore a variety of scenarios without having to collect new data. It can also help in identifying and designing molecules for new drugs or in predicting how different individuals might respond to a treatment based on their genetic or biological characteristics.

It is important to note that although generative AI offers many opportunities, it also raises challenges and ethical concerns, especially regarding the creation of information that could be mistaken for real data or the possibility of using these technologies for deceptive or malicious purposes.

2. Specific applications

Patient selection and establishment of inclusion criteria with generative AI:

Proper patient selection is essential to the success of any clinical trial. Imprecise selection can lead to inconclusive or even erroneous results. Likewise, establishing appropriate inclusion criteria ensures that trial participants are representative of the demographic for which the treatment is intended. Generative AI has the potential to revolutionize these areas in a way that optimizes and refines the selection and judgment process.

1. Prediction of patient profiles: Using generative algorithms, it is possible to model and simulate different patient profiles based on historical and current data sets. These models can predict which types of patients are likely to benefit most from a particular treatment, helping to refine inclusion criteria.

2. Inclusion criteria optimization: Generative AI can explore millions of criteria combinations based on available data. By using simulations, one can determine which combinations of criteria would produce the most relevant and representative group for a specific trial.

3. Identification of patient subgroups: Instead of looking for a single profile, generative AI can identify multiple patient subgroups that could respond similarly to a treatment. This capability is especially useful in clinical trials aimed at treating complex or multifactorial diseases.

4. Synthetic data generation: In situations where data on a specific demographic group is lacking, generative AI can create synthetic data that reflects the characteristics of that group. Although these data are not real, they can offer an approximation that helps researchers design more inclusive trials.

5. Improving the recruitment process: With AI-generated predictions and models, recruitment tools can be developed that identify potential patients more efficiently, reducing the time and effort required to fill participant quotas in a trial.

In conclusion, the application of generative AI in patient selection and establishment of inclusion criteria can lead to more accurate, efficient and representative clinical trials. However, it is essential that researchers use these tools with caution and ethics, always ensuring the safety and well-being of participants.

Design of clinical trial protocols

The design of clinical trial protocols is a meticulous and essential task to ensure that the study is carried out in a systematic, reliable and ethical manner. Traditionally, without the intervention of technologies such as AI, this process has been carried out following a series of well-established steps:

1. Defining the objective: Before anything else, the objective of the essay must be clearly established. It may involve testing the effectiveness of a new medication, comparing two treatments, or evaluating the side effects of an intervention, among others.

2. Choice of study design: There are several types of design, such as randomized controlled trials, observational studies or crossover studies. The choice will depend on the objective of the trial and the particularities of the treatment or intervention under study.

3. Determination of sample size: The number of participants necessary for the study to have statistical validity is decided. This is based on calculations that take into account the expected variability, the desired effect to be detected and the level of statistical significance and power.

4. Establishment of inclusion and exclusion criteria: The characteristics that participants must have to be included in the study and those that would exclude them are defined. These criteria ensure that the appropriate population is selected and minimize risks to participants.

5. Definition of the variables of interest: The variables that will be measured or recorded during the trial are identified, whether they are outcome, exposure or confounding variables.

6. Data analysis planning: Establishes how data will be collected, stored and analyzed. This includes choosing the appropriate statistical techniques to analyze the results.

7. Establishment of monitoring procedures: It is determined how participants will be monitored, how treatments will be administered and how information will be recorded.

8. Ethical considerations: Before starting any trial, it is essential to obtain approval from an ethics committee. Additionally, procedures must be established to ensure informed consent from all participants and to manage any eventuality or adverse effects that may arise during the trial.

This process has traditionally relied heavily on the experience and knowledge of researchers, as well as the exhaustive review of existing literature. Although tested and effective, this method can be laborious, expensive, and subject to human error or bias. The introduction of AI-based tools has the potential to optimize and improve this process at several stages, offering precision, speed and new perspectives in the design of clinical trial protocols.

Incorporating generative AI into the design of protocols for clinical trials has the potential to transform the process, making it more efficient, accurate, and adaptive. Here are some of the ways generative AI can help:

1. Optimization of the study design: Through the analysis and modeling of previous data, generative AI can propose the most appropriate design for a particular trial, taking into account the characteristics of the intervention, the profile of the patients and the variables of interest.

2. Scenario simulation: With generative models, different trial scenarios can be simulated, allowing researchers to anticipate and plan for potential challenges, such as loss of patients to follow-up or the appearance of unexpected adverse effects.

3. Adaptive sample size determination: Generative AI can continuously refine the required sample size based on data as it is collected, ensuring optimal statistical validity while optimizing resources.

4. Design of inclusion/exclusion criteria: Based on models generated from historical and current data, AI can suggest more precise and adaptive criteria, which can improve the

selection of participants and increase the relevance of the results.

5. Variable modeling: Generative AI can identify and model new variables of interest that have not been previously considered, or it can suggest relationships between existing variables that could be relevant to the study.

6. Automation of data analysis: By using generative models, data analysis can be faster and more accurate. AI can detect patterns, relationships or anomalies in collected data more efficiently than traditional methods.

7. Prediction and management of adverse effects: By modeling and simulating patient data, generative AI can help predict the occurrence of adverse effects in specific subgroups of patients, allowing for earlier intervention and management.

8. Continuous protocol improvement: As information is collected during the trial, generative AI can refine and adapt the protocol in real time, improving trial efficiency and ensuring more reliable results.

9. Ethical assessment: Although the final decision on ethical issues must be made by human experts, generative AI can help identify potential areas of concern by simulating different scenarios and their ethical implications.

In summary, integrating generative AI into the design of clinical trial protocols can modernize and optimize the way these studies are conducted. However, it is crucial to maintain a balanced and ethical approach, ensuring that technology is used as a complementary tool and not a substitute for human judgment and experience.

Predicting clinical trial results with Generative AI:

The ability to accurately predict the results of a clinical trial is invaluable. Not only can it save time and resources, but it can also guide clinical and business decisions. Generative AI, with its ability to model and simulate data, offers a powerful tool to anticipate trial outcomes. Below are ways in which generative AI can be applied to predict clinical trial outcomes:

1. Simulation of outcomes based on previous data: Using data from previous trials, generative AI can model and simulate possible outcomes for an ongoing trial. These simulations can provide an early look at the effectiveness or safety of a treatment.

2. Identification of predictive biomarkers: Generative AI can analyze and model complex data sets to identify biomarkers that are related to positive or negative responses to a treatment. This identification can help personalize treatments and improve success rates.

3. Assessment of drug interactions: By simulating data, generative AI can anticipate how different medications will interact when administered together, which can be crucial in predicting side effects or reduced effectiveness of a treatment.

4. Estimation of dropout rates: Using models generated from historical and demographic data, AI can predict which patients are most likely to drop out of the trial, allowing researchers to take preventive measures.

5. Modeling long-term effects: Although clinical trials are often limited in duration, generative AI can use available data to simulate and predict possible long-term effects of a treatment.

6. Real-time adaptation: Based on the data collected during the trial, generative AI can refine its predictions and adapt to new information, offering more accurate projections as the study progresses.

7. Identification of response subgroups: By analyzing and modeling data, generative AI can identify specific subgroups of patients who are more likely to benefit or experience adverse effects, which can influence future clinical decisions.

8. Optimization of resources: By more accurately predicting outcomes, informed decisions can be made about how to allocate resources, whether to continue with subsequent phases of the trial, or whether changes to the study design are necessary.

The use of generative AI in predicting clinical trial outcomes represents an evolution in the way these studies are conducted. However, it is vital that it is used as a complementary tool, combining its predictions with the expertise and knowledge of clinical researchers. By doing so, its potential can be fully exploited, while ensuring the integrity and reliability of the assays.

Analysis and interpretation of results in clinical trials:

Clinical trials generate a vast amount of data, coming from multiple sources and of a heterogeneous nature. From biochemical measurements, medical images, electronic health records, to quality of life questionnaires, each patient brings a multidimensional data set that needs to be processed and analyzed. This information, while rich in content, is also complex and often intricate, presenting unique challenges for analysis and interpretation.

This complexity derives from several factors:

1. Heterogeneity of data: Clinical trials collect information that ranges from quantitative data, such as blood biomarker levels, to qualitative data, such as responses to surveys about well-being or symptoms.

2. Temporality: Data are collected at different time points, from admission to long-term follow-up, which can add a temporal dimension to the analyses.

3. Multiple sources: Data can come from different places, such as laboratories, hospitals, clinics or even patients' wearable devices, each with their own standards and formats.

4. Missing values: It is common for clinical trial data sets to have missing values, either because certain tests were not performed, because patients did not attend some visits, or for other reasons.

5. Interactions and correlations: Data from different sources or types can interact with each other in unanticipated ways, and these interactions can be crucial to interpreting trial results.

Faced with such complexity, traditional analysis and interpretation may be insufficient or inefficient. Conventional statistical techniques, although tested and reliable, are often limited to examining linear relationships or rely on assumptions that may not be appropriate for such complex data. This is where artificial intelligence, and in particular generative AI, can offer advanced tools to address this complexity in a more holistic and precise way.

Given the complex and multidimensional environment of data generated in clinical trials, generative AI is emerging as a valuable tool for uncovering patterns, relationships, and insights that might otherwise be overlooked. The ways in which generative AI can be implemented to enrich the analysis and interpretation of these results are described below:

1. Data reconstruction and complementation: Generative AI, through models such as Generative Adversarial Neural Networks (GANs), can be used to impute missing data, creating synthetic records based on the available information and enriching the data set for a more complete analysis.

2. Discovery of non-linear correlations: Traditional techniques usually focus on linear relationships. However, generative AI can identify and model non-linear correlations between variables, providing a deeper understanding of the underlying mechanisms and interactions.

3. Patient segmentation: Using generative models, it is possible to identify subgroups of patients with similar characteristics or responses, which can be useful for personalizing treatments or identifying specific populations for future trials.

4. Predictive modeling: Based on input data, generative AI can develop models that predict responses to treatments, side effects or any other relevant outcome, improving the ability to anticipate results.

5. Advanced visualization: Generative AI can create visual representations of complex data sets, making it easier to interpret and discover patterns or anomalies.

6. Identification of external factors: Generative models can help discover factors not considered in the initial design of the trial but that influence the results, such as environmental, genetic or lifestyle factors.

7. Dimensionality reduction: For very large data sets, generative AI can help identify the most relevant variables, simplifying analysis without sacrificing accuracy.

8. Cross-validation and verification: By generating synthetic data, simulations and tests can be performed to validate the results of the real analysis, providing an additional layer of confidence in the conclusions.

9. Integration with other data sources: Generative AI can combine clinical trial data with other sources, such as genomic databases or electronic health records, to provide a more holistic analysis.

10. Real-time evolution: As more data is collected during the trial, generative models can be refined and adapted, allowing analysis to evolve over time and adapt to new information.

In conclusion, generative AI, with its ability to model, simulate and analyze complex data sets, offers enormous potential to revolutionize the analysis and interpretation of results in clinical trials. However, it is essential that these tools are used judiciously and in combination with the experience and knowledge of experts in the field, to ensure that the insights generated are accurate and clinically relevant.

3. Conclusions:

The digital revolution and the rise of artificial intelligence have provided a set of powerful tools that have the potential to transform the field of clinical trials. Specifically, generative AI has emerged as a methodology capable of addressing and solving some of the most complex challenges associated with trial planning, execution, analysis, and interpretation.

1. Improved efficiency: Generative AI can optimize multiple stages of the clinical trial, from patient selection to interpretation of results. This optimization can result in a reduction in the time and resources needed to conduct trials, accelerating the availability of innovative treatments to patients.

2. Treatment personalization: The ability of generative AI to identify specific subgroups and model individual responses portends a future where treatments can be more personalized, based on the individual characteristics and responses of each patient.

3. Discoverability: With its ability to identify non-linear correlations, fill in missing data, and model complex interactions, generative AI can uncover insights that might go unnoticed with traditional methods.

4. Adaptability: As more information is collected, generative AI can refine and adapt its models, making it particularly valuable in long-term trials or early-stage research where adaptability is crucial.

5. Data integration: Generative AI's ability to amalgamate and analyze data from diverse sources provides a more holistic and complete view of patients and treatments.

However, despite its countless advantages, it is essential to approach generative AI with caution and a critical perspective. While it can offer valuable insights, its implementation must be supervised by experts to ensure that the results are clinically relevant and ethically sound. Additionally, considerations around data privacy and security are of utmost importance, and must be rigorously addressed.

In short, generative AI presents an exciting and promising horizon for the world of clinical trials. By combining this advanced technology with human experience and knowledge, the sector has the opportunity to make significant advances in the search for more effective and personalized treatments for patients.

References

Garcia-Retuerta D, Canal-Alonso A, Casado-Vara R, Rey AM, Panuccio G, Corchado JM. Bidirectional-Pass Algorithm for Interictal Event Detection. In Practical Applications of Computational Biology & Bioinformatics, 14th International Conference (PACBB 2020). PACBB 2020. Advances in Intelligent Systems and Computing, vol 1240. Springer, Cham. https://doi.org/10.1007/978-3-030-54568-0_20

Castillo Ossa LF, Chamoso P, Arango-López J, Pinto-Santos F, Isaza GA, Santa-Cruz-González C, Ceballos-Marquez A, Hernández G, Corchado JM. A Hybrid Model for COVID-19 Monitoring and Prediction. Electronics. 2021; 10(7):799.
https://doi.org/10.3390/electronics10070799

Intelligent Platform Based on Smart PPE for Safety in Workplaces. Márquez-Sánchez S, Campero-Jurado I, Herrera-Santos J, Rodríguez S, Corchado JM. Sensors (Basel). 2021 Jul 7;21(14):4652
https://doi.org/10.3390/s21144652

A. Canal-Alonso, R. Casado-Vara and J. Manuel Corchado, "An affordable implantable VNS for use in animal research," 2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS), 2020, pp. 1-4,
doi: 10.1109/ICECS49266.2020.9294958

An Agent-Based Clustering Approach for Gene Selection in Gene Expression Microarray. Ramos J, Castellanos-Garzón JA, González-Briones A, de Paz JF, Corchado JM. Interdiscip Sci. 2017 Mar;9(1):1-13
DOI 10.1007/s12539-017-0219-6

Acknowlegments

This study has been funded by the AIR Genomics project (with file number CCTT3/20/SA/0003), through the call 2020 R&D PROJECTS ORIENTED TO THE EXCELLENCE AND COMPETITIVE IMPROVEMENT OF THE CCTT by the Institute of Business Competitiveness of Castilla y León and FEDER funds.