

## Review Article



# Transforming clinical trials: the emerging roles of large language models

Jong-Lyul Ghim <sup>1,2</sup> and Sangzin Ahn <sup>2,3,\*</sup>

<sup>1</sup>Department of Clinical Pharmacology, Inje University Busan Paik Hospital, Busan 47392, Korea

<sup>2</sup>Center for Personalized Precision Medicine of Tuberculosis, Inje University College of Medicine, Busan 47392, Korea

<sup>3</sup>Department of Pharmacology and Pharmacogenomics Research Center, Inje University College of Medicine, Busan 47392, Korea



**Received:** Aug 11, 2023

**Revised:** Sep 12, 2023

**Accepted:** Sep 14, 2023

**Published online:** Sep 19, 2023

**\*Correspondence to**

**Sangzin Ahn**

Department of Pharmacology and  
Pharmacogenomics Research Center,  
Center for Personalized Precision Medicine  
of Tuberculosis, Inje University College of  
Medicine, 75 Bokji-ro, Busanjin-gu, Busan  
47392, Korea.

Email: sangzinahn@inje.ac.kr

**Copyright** © 2023 Translational and Clinical  
Pharmacology

It is identical to the Creative Commons  
Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>).

**ORCID iDs**

Jong-Lyul Ghim

<https://orcid.org/0000-0002-6137-786X>

Sangzin Ahn

<https://orcid.org/0000-0003-2749-0014>

**Funding**

This work was supported by the National  
Research Foundation of Korea (NRF) grant  
funded by the Korean government (MSIT) (No.  
2018R1A5A2021242).

**Conflict of Interest**

- Authors: Nothing to declare
- Reviewers: Nothing to declare
- Editors: Nothing to declare

## ABSTRACT

Clinical trials are essential for medical research, but they often face challenges in matching patients to trials and planning. Large language models (LLMs) offer a promising solution, signaling a transformative shift in the field of clinical trials. This review explores the multifaceted applications of LLMs within clinical trials, focusing on five main areas expected to be implemented in the near future: enhancing patient-trial matching, streamlining clinical trial planning, analyzing free text narratives for coding and classification, assisting in technical writing tasks, and providing cognizant consent via LLM-powered chatbots. While the application of LLMs is promising, it poses challenges such as accuracy validation and legal concerns. The convergence of LLMs with clinical trials has the potential to revolutionize the efficiency of clinical trials, paving the way for innovative methodologies and enhancing patient engagement. However, this development requires careful consideration and investment to overcome potential hurdles.

**Keywords:** Clinical Trial; Natural Language Processing; Artificial Intelligence; Informed Consent; Medical Writing

## INTRODUCTION

Clinical trials, the cornerstone of evidence-based medicine, stand as the gold standard in the field of medical research. These trials necessitate the commitment of substantial resources and the involvement of highly specialized staff with deep expertise in the research process. One of the primary barriers to conducting clinical trials is inadequate funding, often stemming from the labor- and cost-intensive nature of these trials, an issue that becomes even more pronounced as their scale grows [1]. A significant proportion of these trials fail to reach completion, and of those that do, a notable percentage remain unpublished [2]. This leads to potential wastage of both human and financial resources, thereby casting a shadow on efforts to advance medical knowledge.

**Reviewer**

This article was reviewed by peer experts who are not TCP editors.

**Usage of AI tools**

The authors utilized ChatGPT to correct grammar, enhance the readability and generate the summary table.

**Author Contributions**

Conceptualization: Ghim JL, Ahn S; Writing - original draft preparation: Ahn S; Writing - review and editing: Ghim JL, Ahn S.

In recent years, the advent of large language models (LLMs) has ushered in a new era across various industries. LLMs are machine learning models trained on vast amounts of text. They have the capability to predict the subsequent appropriate word (or token) based on a given sequence. By providing a carefully crafted prompt, which is a piece of text that describes the desired objective, the model can generate coherent text sequences that fulfill the given instructions. Through fine-tuning and human feedback, current LLMs have the ability to carry out a multitude of intricate tasks [3]. OpenAI's ChatGPT, Google's Bard, and Anthropic's Claude represent some of the prominent general-purpose LLM-based chatbots accessible to the public. Such accessibility has democratized the capabilities of LLMs, enabling the broader public to harness their power for various tasks. The application of LLMs has been found to exert a disproportionately higher impact on occupations that are highly paid and require extensive training [4]. The application of LLMs to medicine is emerging as a promising and innovative frontier. They have been demonstrated to have considerable potential in handling complex medical information [5]. Within the context of healthcare, LLMs are facilitating revolutionary changes by enabling advanced document generation, the creation of insightful summative reports, and the automation of complex textual output, thus contributing to efficiency and innovation [6].

This review explores the multifaceted applications of LLMs within clinical trials. By examining their most recent applications and thoughtfully considering potential near-future adaptations, we aim to provide a comprehensive and exploratory overview. Our intention is to illuminate the path for researchers, particularly in the domain of clinical pharmacology, to understand and embrace the capabilities of LLMs. In doing so, we seek to encourage wider use of these computational tools in clinical trial practice. The goal of this review is to assist in improving clinical trials and related medical research by addressing the existing gap between technology and healthcare practices.

## ENHANCE PATIENT-TRIAL MATCHING

Patient-trial matching in clinical trials has been an intricate and labor-intensive process, which have attracted the exploration in application of artificial intelligence (AI) [7]. Traditionally, this matching process involves a multi-step approach: meticulous registry screening, comparison of eligibility criteria with the patient's medical profile, and selection of applicable trials based on these assessments. Each stage requires extensive labor performed by skilled personnel, contributing to a time-consuming and often inefficient procedure. Although AI-based trial matching using various algorithms has been explored [8], LLM-based methodologies have recently demonstrated superior adaptability and performance.

Recent innovations have sought to alleviate these challenges through the application of LLMs. In one study, an LLM-based model was used to partially automate the pre-screening process by cross-referencing the medical profiles of candidates with the specific eligibility criteria of different trials [9]. This approach streamlined the evaluation of eligibility criteria, reducing the time and expertise needed for initial screening. Another similar investigation explored a more complex patient-to-trial matching scheme by aggregating criterion-level eligibility data and predicting trial-level eligibility scores [10]. The model's ability to assess and score patients based on multiple factors presented a nuanced understanding of suitability, enhancing the precision of the matching process. Another significant advantage

of LLMs is their generative capability, allowing the model to produce step-by-step reasoning of the output. This transparency enables physicians to review the decision-making process in detail.

It is noteworthy to highlight that the significance of these applications is more apparent in trials involving patients, rather than phase 1 trials recruiting healthy volunteers. To implement these novel approaches, the development of systems to convert electronic health records to patient narratives is essential. These narratives include the overall medical history as well as information needed to compare with eligibility criteria. Such a system can also be equipped with LLMs, which may facilitate this intricate process without requiring structured data [11]. These advancements leverage natural language processing to interpret complex requirements, benefiting both trial-to-patient and patient-to-trial matching. The result is a transformation in clinical trial methodology that enhances efficiency and precision, reduces manual labor, and potentially offers more opportunities to patients.

## STREAMLINE CLINICAL TRIAL PLANNING

Clinical trial planning represents one of the most intricate, time-consuming, and risk-associated aspects of drug development. Traditionally, this process entails the manual curation and analysis of massive amounts of text data, encompassing prior research, regulatory guidelines, and specific therapeutic goals. The synthesis of this information into coherent and compliant trial protocols demands significant expertise and resources.

Recent advancements in LLMs offer promising solutions to these challenges. Their capabilities extend to efficiently processing extensive text data, distilling large volumes of clinical trial descriptions into concise, actionable information. One study exemplified the power of LLMs in this regard, employing an LLM-based model to parse and summarize vast amounts of clinical trial data, thereby aiding investigators in quickly grasping key insights [12].

Beyond summarization, LLMs also possess the ability to generate coherent and context-appropriate text from relatively simple text descriptions. By pretraining on comprehensive corpora of clinical trial documents, researchers have found that LLMs can be harnessed to create criterion descriptions for trials, effectively transforming vague or complex information into clear, standardized language [13].

A more innovative application of LLMs involves predicting the outcomes of clinical trials. Utilizing language models pretrained on a diverse array of clinical trial documents, one study embedded trial results into a mathematical representation, known as embeddings. Analyzing these embeddings according to topics and temporal trends, the researchers were able to generate promising predictions for trial outcomes, paving the way for a more data-driven approach to trial design [14].

These recent advancements reveal a multi-faceted application of LLMs to clinical trial planning. From digesting complex information to generating coherent textual output and even offering predictive insights, LLMs are increasingly seen as valuable tools in enhancing efficiency and mitigating risks. The role of LLMs in clinical trial design signifies a shift towards a more technologically integrated and agile approach. The various applications described above not only simplify existing procedures but also introduce novel

methodologies, all of which contribute to improving the efficiency, accuracy, and innovative potential of clinical trial planning.

## APPLICATIONS ON FREE TEXT NARRATIVES

In the context of clinical trials, data collection is typically structured to facilitate rigorous statistical analyses. Free form text, often a rich source of information, is traditionally coded into predefined criteria to maintain this structured approach. This includes dealing with the old problem of inter-rater reliability in clinical trials [15]. Recent advancements in LLMs have introduced a novel dimension to this process, with the potential to enhance the consistency and accuracy of data coding.

The utilization of domain-specific pre-trained language models, such as those trained on a huge biomedical corpus designed to capture biomedical context-dependent named entity recognition has shown the benefits of the adaptation of LLMs in this area [16]. LLMs possess the ability to label and classify free form text, an attribute that can be integrated into the coding process itself. For instance, studies have demonstrated the feasibility of employing LLMs to automatically classify electronic health records into International Classification of Diseases codes [17]. Furthermore, LLMs have been successfully employed to classify free text in regulatory documents into specific predefined sections [18], and to code text data that require deductive analysis [19], addressing a long-standing challenge in clinical research.

Moreover, the application of LLMs in analyzing patients' free text entries presents an innovative avenue for research. By leveraging the analytical strength of LLMs, it is possible to generate and validate hypotheses concerning differences in text content between groups [20]. This discovery-based methodology has the potential to uncover new insights and possibilities within free text data, an area that may previously have been undervalued or overlooked in clinical trials. The integration of LLMs into this aspect of data analysis may therefore represent a significant contribution to the evolving landscape of clinical research methodologies.

## ASSISTANCE IN TECHNICAL WRITING

Clinical trials consistently require a significant amount of essential documentation. This administrative burden has long necessitated the dedication of substantial time of highly trained staff. However, the emergent application of LLMs to medical document writing offers a promising avenue to mitigate this challenge [6]. The utilization of LLMs to automate diverse forms of paperwork that previously relied on specialized human intervention is gaining traction.

When employing LLMs to draft documents demanding high-level reasoning, the methodology known as “chain-of-thought prompting” might be advantageous. Instead of having the language model to directly output a finalized document, this method guides the model to generate intermediate steps of reasoning, which ensures a coherent and contextually relevant output. The incorporation of this technique has been previously proposed for the automation of property valuation reports, to ensure consistency, objectivity, and transparency in the resulting documents [21]. Within the medical field, LLMs are

being increasingly implemented for routine technical writing tasks. Notable examples of this application are the automated generation of patient discharge summaries, where LLMs synthesize relevant clinical information into a concise and understandable format for patients and caregivers [22]. Another application is the summarization of radiology reports, in which complex imaging findings are distilled into clear and standardized language, making them more accessible to healthcare providers and enhancing the efficiency of diagnostic procedures [23].

Moreover, the capacity of LLMs to proficiently manage and reason from tabular data has been empirically demonstrated [24]. The prospective use of LLMs to compose documents based on provided tables represents a promising avenue to enhance the efficiency of the technical writing process related to clinical trials. This conversion between tabular data and free-form text can work in both directions, furthering the potential applications in technical writing. For demonstration purposes, the manuscript, excluding the introduction and conclusions, was input to ChatGPT and prompted to generate a summary table (**Table 1**). Such automated drafting could substantially alleviate the workload of medical personnel and accelerate the overall workflow of clinical trials, thereby augmenting both the efficiency and efficacy of research processes.

## PROVIDE COGNIZANT CONSENT

The procedure of obtaining informed consent is essential in preserving the rights and safety of participants in clinical trials. Studies have indicated that the comprehension of consent among enrolled patients may be more limited than anticipated [25]. LLMs, often employed as chatbots, enable users to pose questions and receive pertinent information in natural language. Empirical evidence has shown that LLMs can furnish suitable responses to medical inquiries, preserving both the accuracy of information and the empathy in communication [26]. Additionally, LLMs possess the capability to locate and present relevant excerpts of text necessary for responding to specific questions, thereby offering answers based on most recent information to medical queries [27].

By implementing an LLM-powered chatbot equipped with the current trial information, subjects may inquire about the trial and obtain immediate, informed responses. In

**Table 1.** Summary of potential applications of large language models on clinical trials\*

Area of application	Details	Related Examples
Enhance patient-trial matching	Automate pre-screening using LLMs, streamline evaluation of eligibility criteria, and produce step-by-step reasoning of output.	- Cross-referencing medical profiles with eligibility criteria [9]. - Predicting trial-level eligibility scores [10].
Streamline clinical trial planning	Process extensive text data, generate coherent text from simple descriptions, and predict clinical trial outcomes.	- Summarizing clinical trial data [12]. - Creating criterion descriptions [13]. - Predicting trial outcomes [14].
Applications on free text narratives	Enhance the consistency and accuracy of data coding from free text.	- Classifying electronic health records [17]. - Coding text data requiring deductive analysis [19].
Assistance in technical writing	Automate medical document writing and convert between tabular data and free-form text.	- Generation of patient discharge summaries [22]. - Summarization of radiology reports [23].
Provide cognizant consent	Improve comprehension of consent through LLM-powered chatbots and generate text for knowledge gaps.	- LLMs providing answers based on the most recent information [27]. - Assessing knowledge and filling gaps [29].

LLM, large language model.

\*The manuscript, excluding the introduction and conclusions sections, was input into ChatGPT-4.0, and then prompted to create a summary table. For the detailed prompt, refer to the following link: <https://chat.openai.com/share/537912e5-fdb0-481c-aeb2-da1eb29f77da>

educational settings, LLMs have been utilized to automatically generate assessment questions to evaluate students' comprehension [28]. The approach in which LLMs are used to assess current knowledge and fill in the knowledge gaps with dynamically generated text can be adapted for use in clinical trial settings [29]. Such an approach may herald a novel paradigm of “cognizant consent”, actively ensuring comprehension and fostering a more holistic engagement with participants.

## CONCLUSIONS

In this article, the potential applications of LLMs within the context of clinical trials have been examined. Areas of focus include patient screening, clinical trial planning, the analysis of free text narratives, technical writing assistance, and the augmentation of patients' comprehension of trial details (**Table 1**). These innovations may significantly enhance the efficiency of clinical trials, expand the utilization of free text, and contribute to a more robust informed consent process. It must be noted, however, that this review did not utilize a systematic methodology to survey the advancements in LLMs in clinical trials, a factor that could introduce potential biases. Moreover, given the rapid progression in the field of LLMs, many reports are in preprint stages without thorough peer review, and critical studies may have been inadvertently omitted.

The integration of LLMs into clinical practice is not without its challenges, especially from legal and quality assurance perspectives. LLMs are susceptible to generating misleading or incorrect information, a phenomenon known as “hallucination”, and ensuring quality control may prove to be demanding. The complexity and flexibility of LLMs correspondingly make validation regarding accuracy, safety, and clinical efficacy particularly challenging [30]. The inherent opacity of AI models further exacerbates the difficulty of their application in critical, real-world scenarios. Techniques such as chain-of-thought prompting may guide the language model to reveal the reasoning process behind its outputs, thereby increasing the models' explainability [31].

In conclusion, LLMs are catalyzing transformative changes within the medical domain, and clinical trials stand to benefit substantially from these developments. Tasks that are traditionally repetitive and labor-intensive may be conducted more efficiently, while deeper insights may be extracted from free text. The optimization of patient-to-trial matching has the potential to yield benefits for both patients and administrators of trials. The intersection of LLMs and clinical trials merits deliberate investment and scrutiny to expedite this transformative shift.

## REFERENCES

1. Djuricic S, Rath A, Gaber S, Garattini S, Bertele V, Ngwabyt SN, et al. Barriers to the conduct of randomised clinical trials within all disease areas. *Trials* 2017;18:360.  
[PUBMED](#) | [CROSSREF](#)
2. Wang S, Šuster S, Baldwin T, Verspoor K. Predicting publication of clinical trials using structured and unstructured data: model development and validation study. *J Med Internet Res* 2022;24:e38859.  
[PUBMED](#) | [CROSSREF](#)
3. Ouyang L, Wu J, Jiang X, et al. Training language models to follow instructions with human feedback. *arXiv*. March 4, 2022. <https://doi.org/10.48550/arXiv.2203.02155>.  
[CROSSREF](#)



4. Eloundou T, Manning S, Mishkin P, Rock D. GPTs are GPTs: an early look at the labor market impact potential of large language models. arXiv. March 23, 2023. <https://doi.org/10.48550/arXiv.2303.10130>.  
**CROSSREF**
5. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. arXiv. December 26, 2022. <https://doi.org/10.48550/arXiv.2212.13138>.  
**CROSSREF**
6. Atallah SB, Banda NR, Banda A, Roeck NA. How large language models including generative pre-trained transformer (GPT) 3 and 4 will impact medicine and surgery. *Tech Coloproctol* 2023;27:609-614.  
**PUBMED | CROSSREF**
7. Woo M. An AI boost for clinical trials. *Nature* 2019;573:S100-S102.  
**PUBMED | CROSSREF**
8. Idnay B, Dreisbach C, Weng C, Schnell R. A systematic review on natural language processing systems for eligibility prescreening in clinical research. *J Am Med Inform Assoc* 2021;29:197-206.  
**PUBMED | CROSSREF**
9. den Hamer DM, Schoor P, Polak TB, Kapitan D. Improving patient pre-screening for clinical trials: assisting physicians with large language models. arXiv. June 29, 2023. <https://doi.org/10.48550/arXiv.2304.07396>.  
**CROSSREF**
10. Jin Q, Wang Z, Floudas CS, Sun J, Lu Z. Matching patients to clinical trials with large language models. ArXiv. July 28, 2023. <https://doi.org/10.48550/arXiv.2307.15051>.  
**PUBMED | CROSSREF**
11. Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, Parisien C, et al. A large language model for electronic health records. *NPJ Digit Med* 2022;5:194.  
**PUBMED | CROSSREF**
12. White RD, Peng T, Sripitak P, Johansen AR, Snyder M. CliniDigest: a case study in large language model based large-scale summarization of clinical trial descriptions. arXiv. July 31, 2023. <https://doi.org/10.48550/arXiv.2307.14522>.  
**CROSSREF**
13. Wang Z, Xiao C, Sun J. AutoTrial: prompting language models for clinical trial design. arXiv. May 18, 2023. <https://doi.org/10.48550/arXiv.2305.11366>.  
**CROSSREF**
14. Wang Z, Xiao C, Sun J. SPOT: sequential predictive modeling of clinical trial outcome with meta-learning. arXiv. April 7, 2023. <https://doi.org/10.48550/arXiv.2304.05352>.  
**CROSSREF**
15. Berendsen S, Verdegaal LM, van Tricht MJ, Blankers M, Van HL, de Haan L. An old but still burning problem: Inter-rater reliability in clinical trials with antidepressant medication. *J Affect Disord* 2020;276:748-751.  
**PUBMED | CROSSREF**
16. Naseem U, Khushi M, Reddy V, Rajendran S, Razzak I, Kim J. BioALBERT: a simple and effective pre-trained language model for biomedical named entity recognition. In: 2021 International Joint Conference on Neural Networks (IJCNN); July 18-22, 2021; Shenzhen, China. New York (NY): IEEE; 2021, 1-7. doi: **CROSSREF**
17. Huang CW, Tsai SC, Chen YN. PLM-ICD: automatic ICD coding with pretrained language models. arXiv. July 12, 2022. <https://doi.org/10.48550/arXiv.2207.05289>.  
**CROSSREF**
18. Gray M, Xu J, Tong W, Wu L. Classifying free texts into predefined sections using AI in regulatory documents: a case study with drug labeling documents. *Chem Res Toxicol* 2023;36:1290-1299.  
**PUBMED | CROSSREF**
19. Tai RH, Bentley LR, Xia X, Sitt JM, Fankhauser SC, Chicas-Mosier AM, et al. Use of large language models to aid analysis of textual data. *bioRxiv*. July 19, 2023.  
**CROSSREF**
20. Zhong R, Zhang P, Li S, Ahn J, Klein D, Steinhardt J. Goal driven discovery of distributional differences via language descriptions. arXiv. February 28, 2023. <https://doi.org/10.48550/arXiv.2302.14233>.  
**CROSSREF**
21. Cheung KS. Real Estate Insights Unleashing the potential of ChatGPT in property valuation reports: the “Red Book” compliance Chain-of-thought (CoT) prompt engineering. *J Prop Invest Financ*. 2023 [In Press].  
**CROSSREF**
22. Patel SB, Lam K. ChatGPT: the future of discharge summaries? *Lancet Digit Health* 2023;5:e107-e108.  
**PUBMED | CROSSREF**

23. Doshi R, Amin K, Khosla P, Bajaj S, Chheang S, Forman HP. Utilizing large language models to simplify radiology reports: a comparative analysis of ChatGPT3.5, ChatGPT4.0, Google Bard, and Microsoft Bing. medRxiv. June 5, 2023. <https://doi.org/10.1101/2023.06.04.23290786>.  
**CROSSREF**
24. Chen W. Large language models are few(1)-shot table reasoners. arXiv. January 23, 2023. <https://doi.org/10.48550/arXiv.2210.06710>.  
**CROSSREF**
25. Sherlock A, Brownie S. Patients' recollection and understanding of informed consent: a literature review. ANZ J Surg 2014;84:207-210.  
**PUBMED | CROSSREF**
26. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023;183:589-596.  
**PUBMED | CROSSREF**
27. Li Y, Li Z, Zhang K, Dan R, Jiang S, Zhang Y. ChatDoctor: a medical chat model fine-tuned on a large language model meta-AI (LLaMA) using medical domain knowledge. Cureus 2023;15:e40895.  
**CROSSREF**
28. Circi R, Hicks J, Sikali E. Automatic item generation: foundations and machine learning-based approaches for assessments. Front Educ 2023;8:858273.  
**CROSSREF**
29. Ahn S. The impending impacts of large language models on medical education. Korean J Med Educ 2023;35:103-107.  
**PUBMED | CROSSREF**
30. Gilbert S, Harvey H, Melvin T, Vollebregt E, Wicks P. Large language model AI chatbots require approval as medical devices. Nat Med. 2023 [In Press].  
**PUBMED | CROSSREF**
31. Wei J, Wang X, Schuurmans D, et al. Chain of thought prompting elicits reasoning in large language models. arXiv. October 10, 2022. <https://doi.org/10.48550/arXiv.2201.11903>.  
**CROSSREF**