White paper

Harnessing Generative AI for Clinical Trial Documentation: Balancing Innovation with Expert Oversight

Authors

Dayle Cohen, executive director, project management,
PPD clinical research business of Thermo Fisher Scientific

Jonathan Pitt, manager, medical writing, Evidera,
PPD clinical research business of Thermo Fisher Scientific

Sarah Frost, functional lead, FSP medical writing,
PPD clinical research business of Thermo Fisher Scientific

Medical writers play a unique role at the intersection of cutting-edge clinical research and scientific documentation. Whether in pharma, biotech, contract research organizations (CROs), or academic research institutions, medical writers are held to high standards for accuracy, thoroughness, and scientific rigor. This is because the documents medical writers produce are essential for regulatory bodies to make crucial decisions about the efficacy and safety of treatments on behalf of patients. Given the high stakes, changes to tools used in medical writing undergo rigorous evaluation, which can impact the speed of adoption. Those that are adopted often continue to be frequently customized to meet specific organizational requirements and to help ensure that they stand up to regulatory scrutiny.

Employing new approaches, such as generative artificial intelligence (AI) to help automate clinical trial-related writing is a hot topic in the industry. The life sciences industry is increasingly recognizing the potential of AI to reduce the time required to develop regulatory documents, thereby accelerating trials and addressing patient needs more swiftly. These efficiencies are becoming essential as AI disrupts the clinical research space by potentially multiplying the number of molecules entering trials in the near future. This anticipated surge in clinical trials necessitates more streamlined methodologies for drafting and reviewing regulatory and patient-centered documents for final approvals and decisions. Generative AI tools based on large language models (LLMs) are considered well suited to this task, but only when they are overseen by a medical writer who is proficient in the art of prompt engineering, cognizant of the tools' limitations, and equipped with the expertise to verify the accuracy of their outputs.[1]

As pharmaceutical companies and researchers (sponsors) contemplate the integration of generative AI to streamline clinical trial documentation, it is essential to move beyond the initial excitement and critically assess both the capabilities and limitations of these advanced tools. This whitepaper goes beyond the hype to explore the potential of LLM-based generative AI in the creation of clinical and regulatory documents, and how LLMs may impact the medical writing profession. We explore the current landscape, by contrasting the industry's high expectations for AI with its practical limitations, and pinpoint areas where LLMs can provide value and recommend best practices for their optimal use. Crucially, we will underscore the indispensable role of medical writers in ensuring the precision, accuracy, and regulatory compliance of AI-assisted documentation.

Large language models (LLMs), defined. An LLM is a sophisticated machine learning or deep learning model designed to understand, generate, and manipulate human language. These models are trained on extensive datasets encompassing a wide range of text sources, enabling them to perform various language-related tasks such as text completion, translation, summarization, and more. Users typically interact with LLMs through interfaces like chatbots, where they provide prompts and receive coherent and contextually relevant responses.

Current limitations of LLMs for clinical trial content writing.

To understand why LLMs are not the silver bullet for clinical trial content writing, it's important to consider the nature of the content itself. Medical Writing experts at The PPD™ clinical research business of Thermo Fisher Scientific emphasize that the emergent and highly nuanced nature of the information generated through clinical research differs significantly to the volumes of information on which LLMs were initially trained. Here are the key distinctions:

- **Complexity and specificity of clinical research documentation.** Clinical research documents are highly detailed and follow stringent structures and guidelines that evolve over time. The International Council for Harmonization of Technical Requirements of Pharmaceuticals for Human Use (ICH) provides detailed guidelines[2] for clinical study

documentation. Other documents are developed bespoke in response to specific regulatory requests.

- **LLMs are based on historic datasets.** Clinical trials frequently involve new drugs and therapies, which lack extensive pre-existing written information. Foundational LLMs, being trained on large volumes of established content, are not inherently adept at generating clinically relevant content for these emerging therapies.[3] These LLMs also require time to be trained on documents associated with new or changing regulations.

Due to these factors, most foundational LLMs are not suitable as standalone solutions for authoring clinical trial documents, especially those impacted by new or updated regulations. However, by fine-tuning LLMs and designing systems to interface with relevant databases, these models can adapt to the dynamic and nuanced information produced through clinical trials. Also, by leveraging system design principles, LLMs can be engineered to perform reasoning and act on that reasoning (ReAct), thereby improving their performance in tasks such as drafting Informed Consent Forms (ICFs) or reviewing clinical trial protocols. Nevertheless, a significant limitation lies in the necessity for meticulously engineered system architectures to adapt to specific datasets and guidelines, highlighting the importance of ongoing refinement and expert oversight.

PPD Functional Service Partnership (FSP) Medical Writing solutions experts point out that while an LLM can craft easy-to-understand text with efficiency, it is not guaranteed to be free of errors or misrepresentation. For instance, LLMs may detect patterns or objects that do not exist or are imperceptible to human observers, leading to outputs that are nonsensical or entirely inaccurate – a phenomenon known as hallucination. Consequently, there is a risk of producing text that appears plausible but lacks factual accuracy.[4] How can this happen? LLMs are based on statistical probabilities, predicting the next most likely word in a series of words based on a large volume of general-use training data.[5] Their accuracy is reduced with highly specific or specialized contexts.

An article published in the National Library of Medicine concluded that while LLMs can help with background content, their current inability to produce reliable and accurate sources limited their validity as a tool for scientific medical writing.[6] Another study found that GPT-4, a leading LLM, performed relatively poorly at clinical thinking and logic (assessment score just over 40%) compared to other aspects such as correct use of medical terminology.[7]

While the technology is improving, the liability and safety of LLMs in clinical trials pose substantial regulatory, ethical, and practical challenges that are still being addressed by authorities.[7], Therefore, content generated by an LLM must be validated by the skills of scientists, clinicians, and medical writers trained to understand new and complex scientific ideas, especially when errors and misinformation could impact the integrity of the drug development process and the trust of patients who depend on it.[8] LLMs might be best thought of as a starting point rather than a final solution.

Ultimately, the accountability for content lies with the sponsor. To embrace LLMs, there must be a process that allows an expert such as a medical writer to start with LLM-generated content and edit or amend that content to generate a working draft. This process should involve ensuring that the content: a) is scientifically and clinically precise with accurate terminology; b) includes all required elements, key supporting details, and context while omitting irrelevant information; and c) contains appropriate references to literature, regulatory guidelines, and other relevant documents.

PPD FSP Medical Writing solutions experts recommend using LLMs as part of a broader document development process rather than as standalone tools. This process should prioritize quality, accuracy, and sound scientific discussion through ensuring that experts review and refine documents after utilizing LLMs.

Current and future capabilities of large language models.

Now that we've explored the limitations of LLMs, let's analyze how these tools can be most useful in clinical and regulatory documentation. LLMs should be regarded as a helpful assistant or part of a toolkit that supports medical writing. They can be effectively deployed in the following ways, although the user must remain vigilant for potential errors and omissions:

- **Simplifying complexity.** Medical writing often involves highly technical text. LLMs are particularly adept at generating generalized prose for patient-centric documentation, such as informed consent forms and lay summaries, making the content more accessible to the patients and the general public.[9] While medical writers must still verify accuracy, LLMs can significantly reduce drafting time for sections requiring high-level language, such as study backgrounds, indications, or study procedures. This capability is especially beneficial in the European Union, where clear and accessible summaries of clinical trial results for the public are required.[10]
- **Querying complex documents.** LLMs can efficiently summarize the content of complex documents such as Clinical Protocols or Investigator's Brochures. This allows medical writers to quickly understand and query these materials, facilitating a more efficient writing process.
- **Information extraction.** LLMs can be a useful tool for

extracting dense information contained in whitepapers or data-rich tables such as the Schedule of Assessments. This allows the writer to quickly gain context for drafting with greater speed.

- **Supporting administrative tasks.** Medical writers often have numerous administrative responsibilities in addition to the technical writing of the document itself. LLMs can easily support these tasks such as preparing meeting minutes, timelines, and drafting emails, thereby freeing up more time for document authoring.

How could LLMs evolve to augment a medical writer's capabilities?

To truly augment a medical writer's capabilities, the evolution of LLMs should focus on several key areas:

- **Specialized LLM agents:** Developing LLM agents specifically engineered to reliably perform routine tasks could offer significant benefits. Examples include querying study protocols, simplifying the Schedules of Assessments, and converting study results into report-ready descriptions. These specialized agents would streamline processes and enhance the efficiency of medical writers.
- **Enhanced training and user adoption:** Increasing user adoption will require clear and effective LLM training exercises. Medical writers need to be proficient in prompting LLMs to perform specific writing tasks. The quality of the output is directly related to the quality of the prompts. Therefore, medical writers must develop and refine their prompt engineering skills. This will require developing specific standardized prompts for the LLM to extract information from a source document like a study protocol and map it to the placeholders in a standardized ICF template. Understanding how a series of conversational prompts can build context and improve outputs will also be essential. These improvements can help the technology produce more relevant and accurate end products.

Another consideration is how LLMs can be integrated into other technological advances in medical writing. In recent years, one such technology to emerge has been structured content authoring (SCA), which enables users to create modular content components that can be easily retrieved and reused across multiple documents using metadata tags. SCA facilitates the efficient management and reuse of approved content, such as text from protocols, which can be consistently applied across various documents, including study reports, marketing applications, and labels.[11] This not only ensures consistency and compliance but also significantly reduces the time and effort required to produce high-quality documentation.

The integration of AI within SCA tools is further enhancing their capabilities. AI-driven insights from structured data can generate easy-to-understand text, accelerating report creation and improving the overall efficiency of the document development process.[12]

Accelerating AI Integration with Expert Partnerships

To derive the most value from technological advancements, sponsors should prioritize equipping medical writers with a comprehensive toolkit. Medical writers are the primary users of these technologies, and their expertise is crucial in applying the right tools to enhance robust, proven processes. Partnering with technology vendors will create holistic solutions for document development rather than arbitrary automation.

One of the most effective ways to incorporate AI-enabled tools into a sponsor's process is through engaging with an expert partner through an FSP model. Contract research organizations (CROs) offering FSP solutions are often at the forefront of innovation. They actively consult with clients on the latest technologies and best practices, providing ready-to-build solutions tailored to a variety of client needs.

Choosing the right FSP medical writing partner is critical. The ideal partner should offer comprehensive consultation services, not just on a project-specific basis, but also in the selection and development of tools that deliver long-term value. This approach ensures that sponsors can effectively manage the integration and ongoing use of these technologies, which can otherwise be a complex and resource-intensive process.

To maintain the momentum gained from technological innovation, experts with both technical and leadership skills are still needed.

The integration of AI technologies does not diminish the need for human expertise; on the contrary, it amplifies the importance of skilled professionals in ensuring optimal outputs, but most importantly, in ensuring that AI-driven efficiencies do not compromise patient safety and care. We've already highlighted how these technologies are not foolproof and therefore it is essential that experts remain part of the process to identify and correct errors that technology may introduce. But what also will be vital to rapid delivery is that these experts excel in project management, collaboration, and communication.

PPD FSP Medical Writing solutions provide comprehensive support for all facets of medical writing across the development lifecycle. Their experts emphasize that writing is only one component of the skill set required to maximize efficiency. Other critical skills that sponsors should focus on include:

- **End-to-end document project management**
  - Highly skilled medical writing talent is needed to drive

document delivery from inception to completion. Effective project management ensures accountability and mitigates risks related to stakeholder availability, competing demands, and strategic changes. Without these skills, the time saved using technology may be negated.

- **Clinical team management**
  - Medical writers play a pivotal role in engaging clinical teams, promoting efficient and effective collaborative authoring, and facilitating reviews. They ensure that key stakeholders are consulted and that the document meets its intended purpose.
- **Leadership**
  - High retention rates are achieved by partners who focus on strong leadership, employee development, and recognition. Stable resourcing through dedicated teams is crucial, as it retains knowledge of the sponsor's culture and programs, complementing the capabilities of AI-enabled tools.

Conclusion

While LLMs and other technological innovations hold immense potential to transform medical writing, their true value is realized only when combined with the expertise of skilled professionals. Medical writers remain central to this process by ensuring the accuracy, compliance, and quality of regulatory documents. By partnering with medical writing teams that combine the latest technologies with strong leadership and project management skills, sponsors can significantly enhance their regulatory documentation processes.[13,14,15]

About PPD FSP Medical Writing solutions

PPD FSP Medical Writing solutions help biopharmaceutical, biotech, and medical device organizations translate complex scientific data into clear and compliant scientific statements, delivered through the flexibility of an FSP model. For sponsors who are developing and investing in their own technology, PPD FSP Medical Writing solutions provide expert writers who know how to work with, and consult on AI technologies. Learn more about PPD FSP Medical Writing solutions here.

Sources:

1. https://journal.emwa.org/artificial-intelligence-and-machine-learning/ai-language-models-are-transforming-the-medical-writing-space-like-it-or-not/
2. ICH Official web site : ICH
3. https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-024-02459-6
4. Shubo Tian, Qiao Jin, Lana Yeganova, Po-Ting Lai, Qingqing Zhu, Xiuying Chen, Yifan Yang, Qingyu Chen, Won Kim, Donald C Comeau, et al. Opportunities and challenges for chatgpt and large language models in biomedicine and health.*arXiv preprint arXiv:2306.10070*, 2023.
5. https://www.nature.com/articles/s43856-023-00370-1
6. artificial-intelligence-in-clinical-trials-in-the-uk-and-eu-w0410048.pdf (arnoldporter.com)
7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10326069/
8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10547240/
9. https://www.nature.com/articles/s43856-023-00370-1
10. https://health.ec.europa.eu/system/files/2020-02/2017_01_26_summaries_of_ct_results_for_laypersons_0.pdf
11. https://www.fme-us.com/blog/structured-content-authoring-is-transforming-drug-approvals-processes/
12. https://aibusiness.com/automation/yseop-launches-ai-powered-report-automation-platform#close-modal
13. https://www.thelancet.com/journals/landig/article/PIIS2589-7500%2824%2900124-9/fulltext
14. https://hai.stanford.edu/news/large-language-models-healthcare-are-we-there-yet
15. https://www.nature.com/articles/s44184-024-00056-z

Learn more at **ppd.com/fsp-medical-writing**