Use of generative artificial intelligence to create clinical training materials

Valerie Moss,[1] Moamen Hammad,[1] Brian Norman,[1] Wing Yan Yeung,[2] and Patricia Badia Folgado[1]
[1]Prime, London, UK; [2]Bayer Inc., Mississauga, Canada

No. 30

Explore research in more detail by scanning the QR code. A video describing this research can be accessed on the mobile-friendly website version of this poster.

Conclusions

- The quality of the abstracts generated by generative artificial intelligence (genAI) was generally good, capturing the main points and structure of the original papers accurately; however, human supervision was required to ensure the medical details were factually correct, relevant, and that no significant findings were missed.
- GenAI can support with the development of training materials. These were created more rapidly than with a human approach, even when considering additional in-depth strategic validation and fact-checking time.
- As genAI improves over time, this process may be further refined.

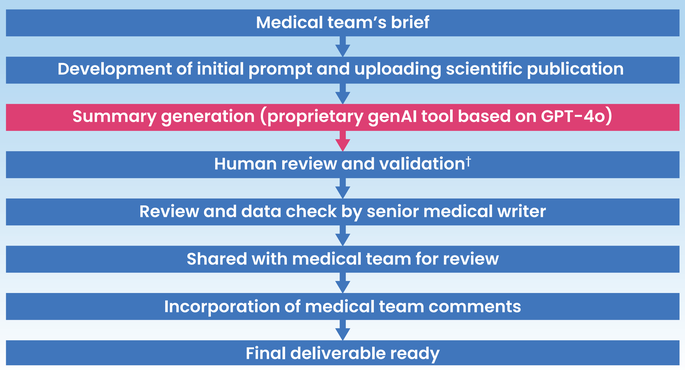
Introduction

- We explored utilization of genAI in generating therapy-area training materials for internal medical pharmaceutical employees.

Methods

- We used a proprietary genAI tool based on GPT-4o to summarize and provide clinical context for seven peer-reviewed publications in a rare cardiology disease.
- The research was undertaken on June 28, 2024.
- An overview of the workflow is shown in **Figure 1**.

Figure 1. Workflow to create clinical training materials using genAI



†A medical writer assessed the summaries for quality and accuracy, making sure the medical details were factually correct, relevant, and not missing any significant findings.
genAI, generative artificial intelligence.

- The genAI was prompted to provide summaries.
  - Feedback was provided to the genAI model to obtain a consistent output.
- The outputs from the genAI were assessed by a medical writer for quality and accuracy, making sure the summaries were factually correct, relevant, and not missing any significant findings.
- This was followed by a review from a senior medical writer and data check.
- Once the writer and senior reviewer were happy with the outputs, they were sent to the client medical director for review and comment.

Acknowledgements: We would like to thank the Synapse Studio and Editor teams at Prime for their support with developing this poster. The authors would like to thank Alexa D'Ambra, of Scion (a division of Prime, London, UK) for her support with this project.

Disclosures: Wing Yan Yeung is an employee of Bayer Inc. and owns shares of Bayer AG. Valerie Moss, Brian Norman, Moamen Hammad, and Patricia Badia Folgado are employees of Prime, London, UK. The research and interpretation provided here represent the views of the authors and not necessarily those of their employers.

Presented at the US Annual Meeting of the International Society for Medical Publications Professionals 2025, Washington DC, US, 11–14 May 2025

Results

- The development of the initial prompt required more than five iterations to reach an output of sufficient consistency and quality.
  - The final prompt used is presented in the box below.

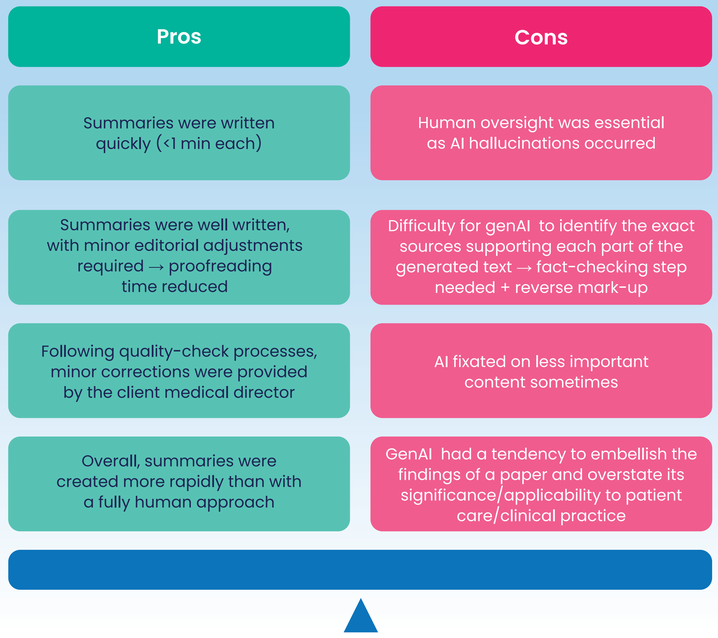
Your role is to summarize scientific publications. The summaries should be up to 800 words for each scientific publication I upload. The scientific publication will be provided as an uploaded document. Follow these instructions:

1. First, include the same title as the one used in the uploaded paper. Literally, exactly the same title.

2. Make the summary appropriate for use as an educational material to onboard new medical colleagues. Every time you summarize a paper, include a title using the same title of the paper in each summary. Include the citation for the paper as well below the title. The summary should have the following headers: 'Introduction/Background information', 'Main findings', and 'Future implications and clinical implications'. The Introduction/Background information should have information about the disease discussed in the paper. Only use information from the uploaded file. Assume a college reading age and include medical information from the paper.

- The main results (pros and cons) are presented in **Figure 2**.

Figure 2. Pros and cons of the methodology



AI, artificial intelligence; genAI, generative artificial intelligence.

- Human oversight was essential as genAI hallucinations occurred.
  - The genAI tool was unable to accurately provide the title and citation of each peer-reviewed publication, despite the prompt being updated five times to explain what was needed.
  - The AI sometimes fixated on less important content for training, such as acknowledgments and references, instead of key findings from the studies.
- The genAI output did not include identification of the exact sources within a paper supporting each part of the generated text, necessitating a comprehensive fact-checking step and a reverse mark-up.
  - Notably, the fact-checker experienced less authority bias when scrutinizing the AI-generated content (**Figure 3**).

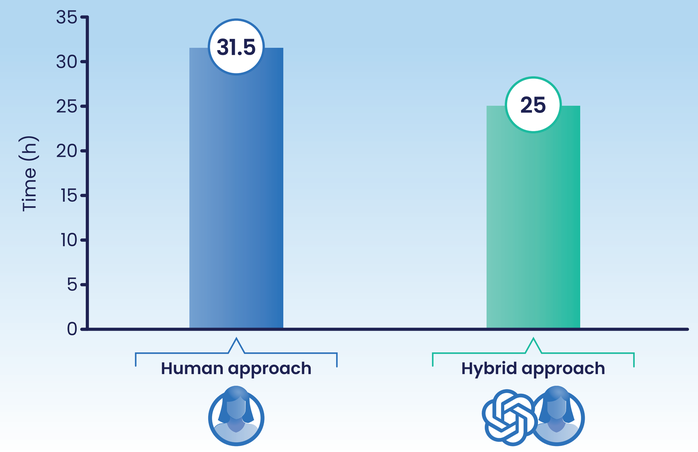
Figure 3. Experience of fact-checking AI-generated content

*"If these were human-written summaries, I think the fact-checking would have taken **less time**."*	→ Fact-checking is believed to have taken less time if content was fully human-developed
*"...there were the **added challenges of determining if the most pertinent details were chosen, if statements belonged in the sections in which they were placed, and if the summaries were consistent across papers**."*	→ The genAI model was unable to consistently identify the most relevant findings
*"The above added challenges led to more comments than I would have otherwise made. This is mostly because if the summaries were written by a person, I may have assumed some of those decisions were intentional and not deemed a comment necessary. To be fair this could be somewhat influenced by..... **my added willingness to critique AI over someone more senior than me**."*	→ Reduced authority bias in the fact-checking step

AI, artificial intelligence; genAI, generative artificial intelligence.

- The estimate of time spent using a traditional human approach versus a hybrid approach is shown in **Figure 4**.

Figure 4. Estimate of time spent using a traditional human approach versus a hybrid approach†



†We estimated that it would take 4.5 hours for each summary if using a traditional fully human approach based on prior experience (including medical writer time, data check, editorial check, and senior review), while the hybrid approach took approximately 3 hours per paper (including the development of the initial prompt).

Limitations

- As this was the first time we had used this genAI model for clinical training materials, there was an element of a learning curve in the iterative prompting. Future projects would benefit from the time invested and may be more time efficient.
- The human time to summarize each paper was an estimate based on prior experience and we did not prepare summaries using a fully human approach as a comparison.
  - Individual expertise, complexity of the papers, and level of analysis required can significantly affect the actual time spent.