

Blinded assessment of potential of Generative Artificial Intelligence (GenAI) to generate plain language summaries

Brian Norman^a, Jennifer Ghith^b and Valerie Moss^a

^aPrime, London, UK; ^bPfizer Inc, Collegeville, PA, USA

In this study, researchers explored if generative AI could help create easy-to-understand summaries of medical research. They used ChatGPT4.0 for this task. ChatGPT4.0 developed two types of summaries: one made by AI alone, and another where a human gave a little more assistance to help ChatGPT make a good summary. These were compared with a summary of the same information written by a human.

A team from a medical communications firm then evaluated each of these summaries. They focused on how well a 16-year-old could understand them. The evaluators did not know who wrote each summary. They found the human-assisted AI summary was as clear as the human-written one.

When asked which summary was the best at being clear and understandable, 44% voted for the human-written one, 36% voted for the human-assisted AI one, and 21% voted for the AI's first attempt.

People who were experts in writing these types of summaries tended to like the human-written one most. However, those new to this field often chose the human-assisted AI version. Science professionals preferred the human-written summary, while people with other job roles preferred the human-assisted AI one.

The study concluded that AI can create clear summaries. However, it performs better with human help and review. The AI summaries were mostly accurate. But, sometimes, they oversimplified things or missed important safety details.

The key takeaway is that AI can aid in writing summaries. Yet, human guidance and accuracy checks are crucial. The study also revealed that opinions on what makes a good summary vary. This depends on a person's experience and profession.

Disclosure: ChatGPT4.0 was used to generate a first draft of this plain language summary. Additional prompts to improve accuracy and improve Flesch-Kincaid Reading Ease score were used, and a human edited the final version.