

# AI's deep research revolution: Transforming biomedical literature analysis

Ying-Mei Wang<sup>a,b,c</sup>, Tzeng-Ji Chen<sup>d,e,f,\*</sup>

<sup>a</sup>Department of Medical Education and Research, Taipei Veterans General Hospital Hsinchu Branch, Hsinchu, Taiwan, ROC;

<sup>b</sup>Department of Pharmacy, Taipei Veterans General Hospital Hsinchu Branch, Hsinchu, Taiwan, ROC; <sup>c</sup>School of Medicine, National Tsing Hua University, Hsinchu, Taiwan, ROC; <sup>d</sup>Department of Family Medicine, Taipei Veterans General Hospital Hsinchu Branch, Hsinchu County, Taiwan, ROC; <sup>e</sup>Department of Family Medicine, Taipei Veterans General Hospital, Taipei, Taiwan, ROC; <sup>f</sup>Department of Post-Baccalaureate Medicine, National Chung Hsing University, Taichung, Taiwan, ROC

Two years ago, shortly after the release of ChatGPT in late November 2022, the *Journal of the Chinese Medical Association (JCMA)* published an editorial, highlighting the impact of ChatGPT and other artificial intelligence (AI) applications on scientific writing.<sup>1</sup> Meanwhile, large language model (LLM) AI has drastically changed the world. The emergence of open source DeepSeek R1 in late January 2025 further shook up the world of AI.<sup>2,3</sup>

The most important feature of the latest development is the advancement of reasoning, which is displayed throughout the iterative processes of thinking about the user's intent, formulating the query, searching the Web or a specific database, reading and summarizing the content, and then responding with a comprehensive report and suggesting follow-up questions. Reasoning is not a unique feature of DeepSeek R1 and has been built into LLMs since their inception. However, a further refinement of the reasoning capability as a specialized AI agent appeared recently. DeepSeek R1 stands out for its transparency ("Explainable Reasoning" to articulate its reasoning process as it arrives at conclusions) and structured problem solving ("Chain-of-Thought Reasoning" to break down complex problems into structured steps), in addition to higher cost effectiveness.<sup>4,5</sup>

The major AI companies have responded quickly by either lowering the prices of existing products or launching similar products, such as Google's Gemini Deep Research (launched on December 11, 2024),<sup>6</sup> OpenAI's Deep Research (February 4, 2025),<sup>7</sup> Perplexity AI's Deep Research (February 14, 2025),<sup>8</sup> and Grok 3's DeepSearch (February 20, 2025).<sup>9</sup> Currently, some of these products are free to try, so that researchers in academia can take advantage of the opportunity to test this new AI capability.

For demonstration purposes, we have tried Grok 3's DeepSearch with a simple query without any hint: "papers about missed appointments since 2023." With this query, we wanted to find recent publications on the topic about which a paper had been published in the *JCMA* 6 years earlier.<sup>10</sup>

Initially, Grok 3 displayed a reasoning framework with a structured eight-step process: thinking, clarifying the request, analyzing the search results, expanding the search scope, combining sources, refining the paper selection, and finalizing the paper list (Fig. 1). It then completed the task in 54 consecutive small steps. The process took 1 minute and 23 seconds.

Grok 3 began by speculating itself about the meaning of the question: missed appointments should mean that a patient misses an appointment at a clinic. Unsure whether the user wanted a list of papers, a summary of abstracts, or specific research findings, Grok 3 decided to find the most recent research on the topic first, trying to find papers in Google Scholar. After Grok 3 found that Google Scholar was not easy to use and there was too much literature, it decided to try PubMed instead, which is suitable for biomedical papers.

Grok3 found 10 journal articles, displayed them one by one, and provided links to the article's journal website, PubMed, or PubMed Central. While considering whether to highlight trends or solutions, Grok 3 found the link to the article content and would try to retrieve relevant information by searching the title and abstracts, findings, and key findings consecutively in the PubMed results.

After discussing the PubMed search results with itself, Grok 3 decided to expand the search scope to the entire Internet and limit the results within Google Scholar. Grok 3 found that the display format was not suitable for retrieving data and then tried to search the Google Scholar site directly, focusing on titles and abstracts. Grok 3 then attempted to combine the PubMed and Google Scholar searches, with 5 or 6 articles in the former and more than 10 in the latter, for a total of 16 articles. Each paper was briefly summarized and the link was provided.

Grok 3 quickly compiled the search results into a summary and a research paper. The summary was structured into key points, overview, detailed findings, and resources. The paper was organized as follows: it began with the title "Comprehensive Analysis of Missed Appointment Research Since 2023," followed by an introduction, background and context, methodology, key findings (which included causes of missed appointments, impacts on healthcare and patients, and interventions and strategies), discussion, and conclusion. The paper also included tables summarizing key studies from PubMed and Google Scholar, categorized by focus area. The main body of the paper, excluding tables and references, consisted of 909 words.

In conclusion, current AI products combine deep thinking and deep search with real-time web search and database access, can focus on peer-reviewed research, offer custom AI workflows, and have their own multimodal capabilities. The citations they find are more transparent and organized, offering paper summaries, evidence extraction, comparison tables, and the ability for

\*Address correspondence: Dr. Tzeng-Ji Chen, Department of Family Medicine, Taipei Veterans General Hospital Hsinchu Branch, 81, Section 1, Zhongfeng Road, Zhudong, Hsinchu 310, Taiwan, ROC. E-mail address: tjchen@vhct.gov.tw (T.-J. Chen).

Conflicts of interest: Dr. Tzeng-Ji Chen, an editorial board member at *Journal of the Chinese Medical Association*, had no role in the peer review process of or decision to publish this article. The authors declare that they have no conflicts of interest related to the subject matter or materials discussed in this article.

*Journal of Chinese Medical Association*. (2025) 88: 415-416.

Received March 18, 2025; accepted March 19, 2025.

doi: 10.1097/JCMA.0000000000001231

Copyright © 2025, the Chinese Medical Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

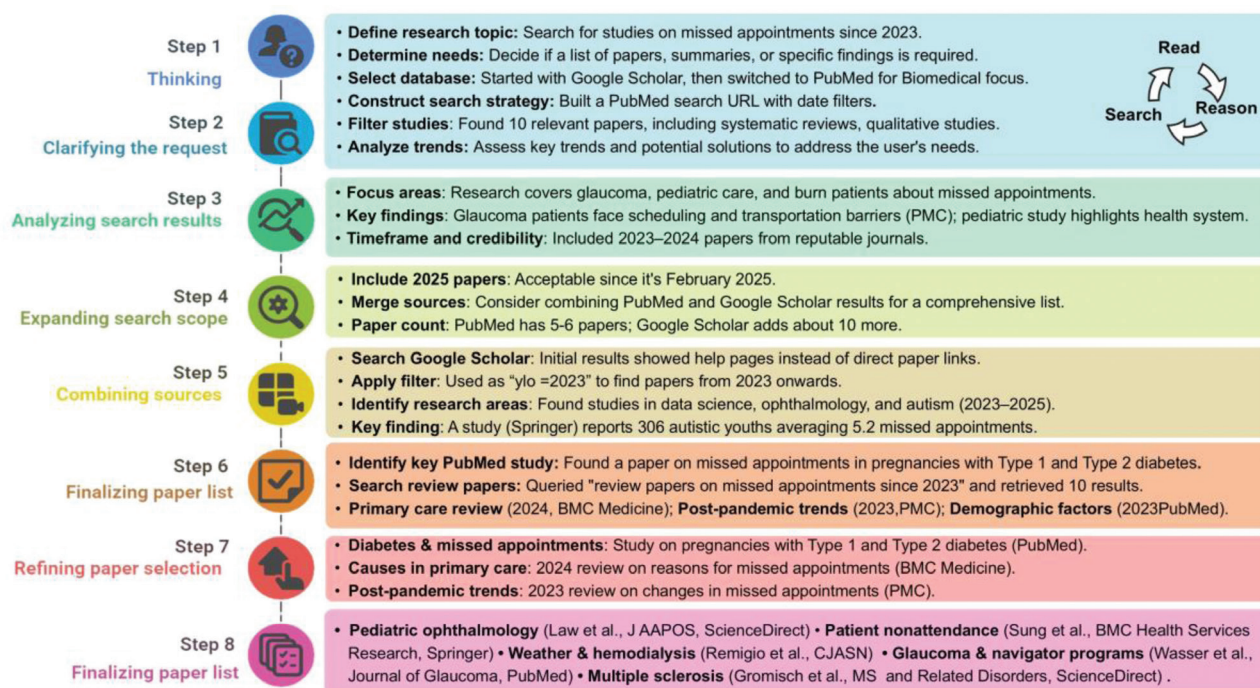


Fig. 1 Workflow of Grok 3's DeepSearch.

users to refine or expand queries on the fly. They also help novices learn academic research logic and enable shareable research reports.

Scientific research can only be compared and reproduced by strictly following fixed methods, which are relatively easy for AI to imitate and learn. Basically, current AI can speed up the retrieval and summarization of literature on one hand, and facilitate the comparison and discussion of results on the other. For dry lab research, especially text-based ones, such as systematic review, meta-analysis, bibliometric study, AI may replace them efficiently. With the dawn of a new era, the future development and face of academia is full of curiosity. Undoubtedly, the crucial part of research will continue to be creative originality and critical thinking.

## REFERENCES

- Chen TJ. ChatGPT and other artificial intelligence applications speed up scientific writing. *J Chin Med Assoc* 2023;86:351–3.
- Conroy G, Mallapaty S. How China created AI model DeepSeek and shocked the world. *Nature* 2025;638:300–1.
- Normile D. Chinese firm's large language model makes a splash. *Science* 2025;387:238.
- Wu J. The rise of DeepSeek: technology calls for the "catfish effect". *J Thorac Dis* 2025;17:1106–8.
- Peng Y, Chen Q, Shih G. DeepSeek is open-access and the next AI disrupter for radiology. *Radiol Adv* 2025;2:umaf009.
- Citron D. Try Deep Research and our new experimental model in Gemini, your AI assistant. Available at <https://blog.google/products/gemini/google-gemini-deep-research/>. Accessed March 18, 2025.
- OpenAI. Introducing Deep Research. Available at <https://openai.com/index/introducing-deep-research/>. Accessed March 18, 2025.
- Perplexity Team. Introducing Perplexity Deep Research. Available at <https://www.perplexity.ai/hub/blog/introducing-perplexity-deep-research>. Accessed March 18, 2025.
- xAI. Grok 3 Beta - The age of reasoning agents. Available at <https://x.ai/news/grok-3>. Accessed March 18, 2025.
- Tsai WC, Lee WC, Chiang SC, Chen YC, Chen TJ. Factors of missed appointments at an academic medical center in Taiwan. *J Chin Med Assoc* 2019;82:436–42.