BMJ Health & Care Informatics

Generative artificial intelligence (AI): a key innovation or just hype in primary care settings?

Annisa Ristya Rahmanti [1,2] Usman Iqbal [3,4,5,6] Sandeep Reddy,[5] Xiaohong W Gao,[2] Huan Xuan Nguyen [2,7] Yu-Chuan Jack Li [8,9,10]

To cite: Rahmanti AR, Iqbal U, Reddy S, *et al*. Generative artificial intelligence (AI): a key innovation or just hype in primary care settings? *BMJ Health Care Inform* 2024;**31**:e101367. doi:10.1136/bmjhci-2024-101367

Received 12 November 2024
Accepted 17 December 2024

For numbered affiliations see end of article.

Correspondence to
Prof Yu-Chuan Jack Li; jack@tmu.edu.tw

The rapid development of generative artificial intelligence (GenAI), particularly large language models (LLMs) like GPT-4, is reshaping healthcare. Built on transformer architectures and trained with extensive datasets, these models are not only accurate but also highly adaptable to various healthcare applications.[1] In the medical field, LLMs are showing promise in automating diagnostic support, improving patient therapeutic management and facilitating clinical decision-making processes.[2–4] These applications are transforming the efficiency and quality of care across patient care delivery.

As the healthcare landscape evolves, a key question arises: How quickly are these innovations being adopted, and by whom? A study by Blease *et al*[5] reveals that UK general practitioners (GPs) are already integrating AI technologies into their practices, even without formal guidelines. This early adoption indicates a shift in the healthcare sector, with clinicians using GenAI to enhance efficiency and decision-making, despite the challenges that come with it.

One major benefit that GPs report is the ability of GenAI to streamline administrative tasks, particularly by generating documentation after patient appointments. According to the survey, 29% of GPs use GenAI for this purpose, saving time and allowing more focus more on patient care. Similarly, a study found that 76.9% of physicians felt more efficient using GPT-4 to draft patient messages, further suggesting that AI can reduce time spent on administrative tasks and improve consistency.[3]

Beyond administrative functions, GPs also use GenAI to assist with clinical challenges. For instance, 28% of GPs use it for differential diagnosis by analysing large volumes of clinical data and suggesting potential diagnoses, potentially reducing diagnostic errors. GenAI's success in this area is partly due to Chain of Thought (CoT) reasoning, which allows the model to break down complex decision-making processes into logical steps, making suggestions more interpretable for clinicians.[6] Additionally, 25% of GPs use GenAI to suggest evidence-based treatment options, enhancing clinical decision-making, especially in complex cases.[5]

CHALLENGES AND RISKS

Despite the potential benefits, integrating GenAI into clinical practice presents several challenges. One major concern is the risk of 'AI hallucinations', where the GenAI generates inaccurate or misleading information.[1] Over-reliance on AI outputs can undermine clinical judgement, potentially leading to diagnostic or treatment errors.[3] While GenAI can process vast amounts of data quickly, it cannot replace the critical thinking and expertise that clinicians provide. Balancing GenAI support with clinical oversight is essential to ensure patient safety.

Another challenge is algorithmic bias.[7] AI models like GPT-4, though valuable for clinical decision support, can perpetuate racial and gender biases across healthcare tasks. A study by Zack showed significant bias in GPT-4's clinical vignettes, differential diagnoses and treatment recommendations.[7] To mitigate this, AI models must be developed and fine-tuned with diverse datasets that represent different demographic groups, promoting equity in healthcare delivery.[8]

Patient privacy and data security also major concerns. Many GenAI tools used in healthcare are driven by private, for-profit companies, raising questions about the handling and protection of sensitive health data.[9] Without clear privacy regulations and transparent data policies, patient trust in AI-driven healthcare systems could erode, leading clinicians to be hesitant about adopting these technologies.

Finally, there is concern over how GenAI is used inconsistently across clinical practice. Some GPs may over-rely on AI tools, while others might avoid them due to uncertainty about their reliability. To address this, validation studies—including expert reviews, clinician-led trials and patient-centred feedback—should be prioritised. These studies will help build the evidence base needed to optimise GenAI use and ensure its safe integration into real-world clinical settings.

FUTURE DIRECTIONS

While the study by Blease *et al* offers valuable insights, it also has limitations that restrict the broader applicability of the findings. These limitations underscore the need for more comprehensive, real-world data. The study also highlights key risks but falls short of providing actionable recommendations to mitigate them. Future research should move beyond merely identifying risks and focus on developing frameworks that ensure the safe and ethical use of GenAI in healthcare.[10] This includes training clinicians to effectively use GenAI as a support tool and conducting long-term evaluations of its impact on clinical outcomes and healthcare disparities. Further, regulatory authorities need to establish frameworks that incorporate principles of transparency, data security, risk management and clinical evaluation metrics to ensure any GenAI applications are safe to use and effective.[11]

Unsafe LLM-based applications, as noted by Blease *et al*, are already entering the market with limited validation. Without careful implementation, there is a risk that GenAI could be misused or relied on improperly. Ensuring the safe integration of GenAI into healthcare requires not only technical refinement but also ethical oversight and evidence-based practices. With responsible implementation, GenAI has the potential to revolutionise healthcare delivery, improving clinical outcomes and patient safety.[1]

Author affiliations
[1]Department of Health Policy and Management, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia
[2]Faculty of Science and Technology, Middlesex University London, London, UK
[3]School of Population Health, Faculty of Medicine and Health, University of New South Wales, Sydney, New South Wales, Australia
[4]Department of Public Health and Community Medicine, Sheikh Zayed Postgraduate Medical Institute, Lahore, Pakistan
[5]School of Public Health and Social Work, Faculty of Health, Queensland University of Technology, Brisbane, Queensland, Australia
[6]Public Health Department, Faculty of Medicine, Universitas Negeri Semarang, Semarang, Indonesia
[7]International School, Vietnam National University, Hanoi, Vietnam
[8]Graduate Institute of Biomedical Informatics, College of Medical Science & Technology, Taipei Medical University, Taipei, Taiwan
[9]Department of Dermatology, Taipei Municipal Wanfang Hospital, Taipei Medical University, Taipei, Taiwan
[10]TMU Research Center for Artificial Intelligence in Medicine and Health, Taipei, Taiwan

Contributors ARR: conceptualisation, writing-original draft. UI: conceptualisation, supervision, writing-review and editing. SR: supervision, writing-review and editing. XWG: supervision, writing-review and editing. HXN: supervision, writing-review and editing. YCJL: conceptualisation, supervision, writing-review and editing. YCJL is the guarantor.

Funding ARR, XG and HXN are supported by the British Council Women in STEM Project Number WISF22-005, the British Council ISPF Research Collaborations Programme, Grant ID # 1203701580, and the Academy of Medical Sciences' ISPF Networking Grant, Grant ID: NGR1\1961.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; internally peer-reviewed.

Author note During the preparation of this work the authors used ChatGPT for language editing. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

ORCID iDs
Annisa Ristya Rahmanti http://orcid.org/0000-0001-9478-6267
Usman Iqbal http://orcid.org/0000-0002-0614-123X
Huan Xuan Nguyen http://orcid.org/0000-0002-4105-2558
Yu-Chuan Jack Li http://orcid.org/0000-0001-6497-4232

REFERENCES

1 Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *NPJ Digit Med* 2023;6:120.
2 Iqbal U, Lee LT-J, Rahmanti AR, *et al*. Can large language models provide secondary reliable opinion on treatment options for dermatological diseases? *J Am Med Inform Assoc* 2024;31:1341–7.
3 Chen S, Guevara M, Moningi S, *et al*. The effect of using a large language model to respond to patient messages. *Lancet Digit Health* 2024;6:e379–81.
4 Lu MY, Chen B, Williamson DFK, *et al*. A multimodal generative AI copilot for human pathology. *Nature New Biol* 2024;634:466–73.
5 Blease CR, Locher C, Gaab J, *et al*. Generative artificial intelligence in primary care: an online survey of UK general practitioners. *BMJ Health Care Inform* 2024;31:e101102.
6 Lee G-G, Latif E, Wu X, *et al*. Applying large language models and chain-of-thought for automatic scoring. *Comput Educ Artif Intell* 2024;6:100213.
7 Zack T, Lehman E, Suzgun M, *et al*. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. *Lancet Digit Health* 2024;6:e12–22.
8 Hastings J. Preventing harm from non-conscious bias in medical generative AI. *Lancet Digit Health* 2024;6:e2–3.
9 Elyoseph Z, Gur T, Haber Y, *et al*. An Ethical Perspective on the Democratization of Mental Health With Generative AI. *JMIR Ment Health* 2024;11:e58011.
10 Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare (Basel)* 2023;11:887.
11 Reddy S. Global harmonisation of AI-enabled software as a medical device regulation: addressing challenges and unifying standards. *SSRN* 2024.