Invited Commentary | Health Informatics

A Patient-in-the-Loop Approach to Artificial Intelligence in Medicine

Maxime F. Griot, MSc, MD; Graham A. Walker, MD

Author affiliations and article information are listed at the end of this article.

The rapid development of artificial intelligence (AI) in medicine has far outpaced its clinical practice implementation. This gap stems largely from the high safety and trustworthiness standards expected and required by health care practitioners.[1] Patient acceptance remains a substantial hurdle, as many lack sufficient information to form opinions on AI; in some cases, this uncertainty even leads to resistance against the integration of AI in their care.[2]

A large-scale survey by Busch et al[3] examined patient perceptions of AI in health care, gathering responses from 13 806 patients across 43 countries in 2023. The study found that while approximately half of respondents held favorable views of AI, acceptance was higher among those familiar with AI or those with higher technological literacy as measured by the number of technology devices used weekly.

One of the key insights of the study by Busch et al[3] is the divergence in AI perception based on patients' health status. Patients in poorer health were more likely to view AI negatively compared with those in better health. (While this finding is compelling, the study did not control for confounding variables, such as general attitudes toward health care.) We agree with the authors, who discuss 2 plausible causes: the influence of depression associated with chronic illness, and the negative experiences accumulated through frequent health care interactions.[4] We would further add that perceived autonomy may also play a role. More medically at-risk patients may feel less in control of their care and have specific concerns about computers further diminishing their agency in decision-making that directly affects their health.

Current research on AI in medicine predominantly focuses on evaluating model performance in specific tasks, such as interpreting imaging studies, summarizing clinical notes, or diagnosing from vignettes.[5] While these studies sometimes demonstrate the superiority of AI over physicians in isolated tasks, more realistic scenarios, such as patient history-taking, reveal substantial limitations.[6] These mixed results suggest that the role of AI is best envisioned as augmenting, rather than replacing, physicians' care processes. This notion aligns with the survey findings, in which only 4.4% of patients favored the use of fully autonomous AI systems without physician oversight.[3]

Furthermore, the survey findings by Busch et al[3] revealed that patients prefer less accurate AI systems that are explainable over highly accurate but opaque "closed box" models. This preference underscores the importance of transparency and the ability to audit AI-assisted processes, reinforcing the notion that patients value augmented health care processes in which physicians maintain oversight. Interestingly, the emphasis of patients on explainable AI aligns with the perspective of physicians, as many express a strong intention to integrate AI into their practice but stress the need for transparency, trust, and explainability in these tools.[7]

Involving patients in the design and development of AI systems can play a pivotal role in making these technologies more acceptable and aligned with patient values. This approach reduces the risk of a mismatch between the capabilities of deployed solutions and the expectations of patients, ensuring that the systems are tailored to clinical practice needs. Moreover, active patient involvement fosters trust, a critical factor for the adoption of AI in health care. Combining this with a physician-in-the-loop strategy could provide a robust framework for AI integration, with both patients and clinicians shaping the development and use of these tools. Rather than aiming to outperform clinicians on narrowly defined tasks, this collaborative approach emphasizes augmenting care processes in ways that resonate with all stakeholders involved.

In conclusion, general attitudes about AI in health care from a broad survey of patients in 43 countries in 2023 were more positive than negative, but not overwhelmingly so. There was strong survey agreement around AI augmenting and improving the capabilities of a patient's physician, and concern about physicians being outright replaced by AI solutions. It is also clear that there were substantial variations in patients' opinions on AI, and this study highlighted interesting differences in attitudes, especially depending on a patient's relative knowledge about AI, their medical complexity, and health status. As more people are exposed to AI tools and algorithms, one may expect these opinions to change over time, since more people will become more knowledgeable and more comfortable with them in their daily lives. We believe an approach that fundamentally centers AI tools around the patient-physician relationship, with an AI tool that can help physicians but not replace them, is the approach that aligns with the values that are suggested from this study by Busch et al.[3]

Corresponding Author: Maxime F. Griot, MSc, MD, Institute of NeuroScience, Université catholique de Louvain, Avenue Mounier 53, boîte B1.53.02, 1200, Woluwe-Saint-Lambert, Belgium (maxime.griot@uclouvain.be).

Author Affiliations: Institute of NeuroScience, Université catholique de Louvain, Brussels, Belgium (Griot); Louvain Research Institute in Management and Organizations, Université catholique de Louvain, Louvain-la-Neuve, Belgium (Griot); Department of Emergency Medicine, Kaiser San Francisco Medical Center, San Francisco, California (Walker).

REFERENCES

1. Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: scoping review. *JMIR Hum Factors*. 2024;11(1):e48633. doi:10.2196/48633

2. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. *J Med Internet Res*. 2021;23(11):e25856. doi:10.2196/25856

3. Busch F, Hoffmann L, Xu L, et al; COMFORT Consortium. Multinational attitudes toward AI in health care and diagnostics among hospital patients. *JAMA Netw Open*. 2025;8(6):e2514452. doi:10.1001/jamanetworkopen.2025.14452

4. Eriksen AA, Fegran L, Fredwall TE, Larsen IB. Patients' negative experiences with health care settings brought to light by formal complaints: a qualitative metasynthesis. *J Clin Nurs*. 2023;32(17-18):5816-5835. doi:10.1111/jocn.16704

5. Bedi S, Liu Y, Orr-Ewing L, et al. A systematic review of testing and evaluation of healthcare applications of large language models (LLMs). *medRxiv*. Preprint posted online August 16, 2024. doi:10.1101/2024.04.15.24305869

6. Johri S, Jeong J, Tran BA, et al. An evaluation framework for clinical use of large language models in patient interaction tasks. *Nat Med*. 2025;(31):77-86. doi:10.1038/s41591-024-03328-5

7. Liu CF, Chen ZC, Kuo SC, Lin TC. Does AI explainability affect physicians' intention to use AI? *Int J Med Inform*. 2022;168:104884. doi:10.1016/j.ijmedinf.2022.104884