1 Uppsala University, Sweden
2 Digital Psychiatry, Beth Israel Deaconess Medical Center, Boston, USA
3 Abbey Medical Centre, London, UK
4 Macquarie University, Sydney, Australia
5 Comunidad de Madrid Servicio Madrileno de Salud, Spain
6 King's College London, London, UK
Correspondence to: C Blease
charlotte.blease@uu.se
Cite this as: *BMJ* 2025;391:e085323
http://doi.org/10.1136/bmj-2025-085323

How generative AI affects patient agency

Generative AI is expanding access, understanding, and choice for patients but also creating new risks and ethical questions, argue Charlotte Blease and colleagues

Charlotte Blease, [1,2] Marcus Lewis, [3] Sara Riggare, [1] David Fraile Navarro, [4,5] Richard Lehman[6]

What once sounded like science fiction is now a real health encounter. Take a recent case that went viral on social media.[1] A patient receives a serious diagnosis from their doctor. Concerned, they input details of the medical scan into the artificial intelligence (AI) platform ChatGPT, which offers a less alarming interpretation. When the doctor re-examines the case, they admit the AI was correct and acknowledge the error.

This and many other patient stories[2] raise urgent questions about how generative AI is shaping clinical care. Public facing generative AI tools have been available for only two years, and rigorous, peer reviewed analysis on their uptake and impact from patients' perspectives is lacking. Nonetheless, the rapid advancement of large language models (LLMs) generative AI could profoundly alter the patient experience, giving rise to triadic care, in which AI is used and openly discussed in the clinical encounter.[3]

Consumer applications, such as OpenAI's ChatGPT, Google's Gemini, and DeepSeek, can provide patients with unprecedented access to medical information, the potential for diagnostic insights, and even offer compassionate support.[4,5] They are changing how patients access and act on health information, but unless both autonomy and agency are explicitly supported, these tools may exacerbate rather than reduce existing inequalities. Autonomy refers to patients' ability to make informed and meaningful choices; agency concerns their capacity and support to act on those choices. Strengthening one without the other risks creating a gap: patients may be offered rights they are unable to exercise or enabled to act without having truly chosen. Although generative AI applications span multiple specialties, this article, part of a BMJ series on generative AI in the clinical encounter (www.bmj.com/collections/gen-AI), focuses primarily on their implications for patients in primary care, including how the growing use of ambient AI scribes for real time documentation influences the consultation experience.

Generative AI as gateway to medical expertise

The patient journey into care is evolving with access to the internet and new generative AI tools (fig 1). Despite limited evidence on how patients are currently using generative AI, these tools may offer new avenues for improving access to care, which includes not only the availability of services but also their affordability, accessibility, and acceptability. For patients in under-resourced areas, such as rural communities, and those who rely on public transport or who face long wait times or high costs, generative AI may offer low cost, on-demand access to health information, even if it does not yet replace input from clinicians. These tools may also improve accessibility by simplifying medical language, offering multilingual support, or being more easily usable by patients with disabilities or limited health literacy.[6] In terms of acceptability, there is evidence that generative AI may reduce stigma, particularly in mental health, by offering anonymous support to those who feel uncomfortable discussing their concerns with a clinician.[7]

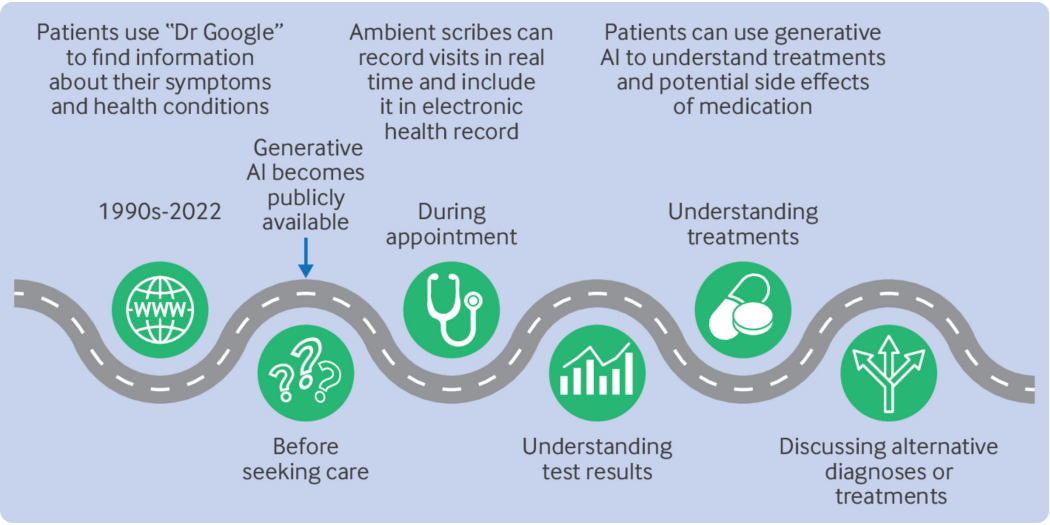

Fig 1 | Evolution of the patient journey with access to AI

Despite the potential of AI for widening access to expertise, uptake and benefit are unlikely to be evenly distributed. Younger people are likely to be more inclined to adopt these tools[8] whereas patients already affected by digital inequities, such as older adults and people in low income households or living in areas with limited connectivity, may struggle to access or effectively use AI powered tools. These groups are particularly vulnerable to being left behind, exacerbating existing inequalities in access to healthcare.

Access is empowering only if it is genuinely accessible—linguistically, culturally, and contextually. While advances in generative AI have improved access for people who speak majority languages,[9] those who speak less widely used languages face continued exclusion.[10] For example, the terminology and framing adopted by generative AI health tools often reflect US centric communication norms and healthcare system structures, which may undermine their cultural relevance and resonance for patients in different healthcare models and social contexts.[6] Moreover, access to information alone may be insufficient if it is not accompanied by access to essential treatments, medications, or even basic infrastructure such as clean water and sanitation. In such contexts, information can deepen frustration or reinforce access inequities, rather than reduce them.[11]

Patient autonomy

Autonomy is the right and capacity of patients to make informed choices that reflect their values, free from undue external control. Unlike agency, which is about the power to act, autonomy emphasises self-determined decision making. Therefore, generative AI's role in supporting patient autonomy depends on how it mediates decision making. Support for patient autonomy requires more than the passive delivery of information; it involves enabling patients to interpret, weigh, and act on that information in ways that reflect their values and circumstances.

Studies suggest that generative AI tools can improve patient understanding by simplifying medical language and tailoring content to individual needs, enhancing clarity and engagement in care decisions.[12] Patients often struggle to understand or remember what their doctor is telling them during time constrained medical visits. Information from generative AI can be accessed as required and may be more comprehensive than that delivered in a consultation.[13]

However, generative AI outputs are not inherently neutral—they reflect underlying assumptions and framing that can subtly shape patient decisions.[11] Because these models are often trained on non-representative data, their recommendations can systematically underserve or mischaracterise the needs of patients from marginalised communities.[14] Further research is needed to determine when, how, and for whom generative AI improves comprehension, particularly compared with information given by clinicians, and how it improves and aids clinicians in delivering health information.

Relatedly, autonomy isn't just about having a choice; it is about being meaningfully able to exercise that choice, with support and recognition of your own circumstances. Patients' ability to make decisions based on AI information is shaped by how they perceive the tools: their cultural fit and perceived authority, and whether they reinforce or reduce past experiences of exclusion. While some recent studies suggest that generative AI can improve the clarity and readability of medical information, most evidence remains experimental and is often based on clinicians rating AI generated content in blinded scenarios[5] rather than patient perspectives.

AI scribes may also improve the accuracy of records and free clinicians' time for patient discussion, indirectly supporting agency. However, they raise privacy and consent concerns, including how patients are informed, how recordings are stored, and how data may be used.[15] Moreover, because AI scribes determine what is recorded in the patient's medical record—and thus what patients later see and interpret—they can indirectly shape patient understanding, autonomy, and trust. Finally, generative AI's tendency for sycophantic responses means it may not present the most appropriate options—for example, tailoring lifestyle advice to echo the user's stated preferences without presenting evidence based alternatives. Such superficial "choices" that mirror algorithmic biases rather than genuine patient priorities risk eroding true autonomy.[16]

Added to this is the concern about the accuracy of patient information. Generative AI tools are renowned for making authoritative sounding errors ("hallucinations"), and a high proportion of their recommendations can be dangerously wrong in health contexts.[17] For example, a recent case described a man who developed bromism after following ChatGPT's dietary advice to replace table salt with sodium bromide, highlighting how decontextualised AI guidance can mislead patients and cause harm.[18] These inaccuracies are not evenly distributed: evidence shows that biases in training data can lead to worse recommendations for women and minoritised groups for many conditions, reinforcing existing health inequalities.[19] Despite these challenges, generative AI tools may outperform clinicians in certain contexts, particularly in rare diseases, where medical professionals often lack knowledge and patients report receiving inaccurate, delayed, or outdated information.[20]

Turning patient choices into action

Generative AI tools have the potential to enhance patient agency by offering personalised, accessible health guidance outside the clinical encounter. This may be particularly valuable for patients who feel uncertain, self-conscious, or fear being a burden when raising concerns with clinicians. AI's ability to provide instant, round-the-clock responses may offer patients a sense of support, particularly when processing difficult news or when they have inadequate access to health professionals.[7]

Independent access to health information may also be valuable, particularly for patients who feel dismissed in clinical settings or excluded from traditional forms of expertise. However, genuine agency depends not just on having a tool, but on being equipped to use it in ways that are meaningful, safe, and aligned with your needs. For example, a generative AI tool that helps schedule medication reminders or find accessible exercise programmes can bolster agency by turning decisions into actionable steps.

Similarly, there is evidence that patients managing rare diseases or chronic conditions are already using generative AI tools such as ChatGPT to fill persistent information gaps from conventional care. For example, a patient with a rare condition could use a generative AI tool to better understand test results and prepare questions for an upcoming consultation, which could help them engage more confidently in shared decision making. In 2025, a US survey found that adults who live in households that have someone with a rare disease were more likely than other households to engage in digital health practices, including telehealth consultations, online peer support, and using AI for health inquiries.[20]

Yet, agency is not equally distributed. As well as those with low digital literacy and reduced access to digital tools, people who lack dexterity, have speech impediments, or face cognitive challenges

may find it difficult to use generative AI. Furthermore, the touted interactivity of "digital first" health services often creates a misleading veneer of patient control; in reality, opaque navigation structures and concealed default settings subtly nudge people towards predetermined actions in which commercial imperatives, such as keeping people interacting with chatbots and disclosing sensitive health information, routinely over-ride genuine patient engagement.

Finally, patient trust in generative AI is essential for true agency: patients will act on its recommendations only if they deem them reliable, secure, and free from bias. Without that trust, even the most personalised guidance risks being ignored or second guessed (including by doctors), undermining patients' capacity to translate choice into action (box 1). Equally, driving cultural change among clinicians could foster trust in generative AI tools, lessening administrative burdens and thereby giving clinicians more time to strengthen patient agency.

Box 1: Patient trust in AI

Public confidence in AI for healthcare is low. A Pew Research study found that 60% of Americans would feel uncomfortable if their healthcare provider relied on AI in their own medical care.[21] A UK-wide survey found that although the public broadly supports the use of AI to assist doctors, especially in radiology, most are uncomfortable with AI replacing human clinicians.[22]

Nevertheless, as AI assisted diagnosis and management plans become more familiar, patients may begin to view generative AI integration as part of standard care. Transparent and safe use will be essential to build trust.

Data transparency and informed consent

- Data handling must be transparent—patients have a right to know how their information is collected, stored, retained, shared, and deleted
- Plain language informed consent is essential to uphold autonomy and maintain public trust
- Privacy standards should be enforced at institutional, regional, and national levels to prevent misuse, including unauthorised tracking or cross border commercialisation of health data
- Clinicians and institutions have a duty to ensure meaningful consent for emerging technologies, such as the expanding use of ambient AI scribes

Embedding trust, equity, and safety in generative AI's future

To fully support patients, generative AI interventions must both respect autonomy by offering unbiased information aligned with their values and foster agency by equipping patients with actionable pathways and support structures to follow through on their decisions. Robust regulation, education, and ethical safeguards are essential to ensure AI enhances healthcare without compromising patient safety. Policy makers must enforce transparency, independent validation, and accountability in AI driven health tools while strengthening digital health literacy to help patients critically assess AI generated advice.

Equally, strict data protection measures must prevent the exploitation of health information, ensure ethical AI development, and maintain patient trust. Even well intentioned AI can introduce gaps in accountability and trust. Errors or opaque reasoning in AI recommendations may leave patients and clinicians unsure who is responsible for adverse outcomes, while reports of misdiagnosis or over-testing can erode confidence in both the tool and the treating physician (box 2).

Box 2: Policy responses to challenges of AI

Improve accuracy in patient facing AI tools

Regulators should require all patient facing AI tools promoted as "medical grade" topass a standardised "health advice accuracy" test before market approval and mandate regular independent audits of outputs with public reporting of errors. The EU AI Act's focus on transparency and safety in healthcare offers new reference points for these standards. In addition, Spain has introduced strict labelling laws for generative AI content, including in health contexts, with penalties for non-compliance.[23]

Ensure AI benefits the global majority

Governments and health authorities should follow the lead of India, which has embedded regional language support and mobile access in its digital health tools.[24] Local and federal governments must mandate use of local patient data—or high fidelity proxies—for training AI tools and require annual cultural and language reviews together with academic and community oversight.

Strengthen digital health literacy

The NHS is rolling out AI literacy programmes for clinicians and patients to increase awareness of the risks of relying on AI generated diagnoses. Accredited AI bias and safety modules should become part of clinician revalidation, and government should fund community "AI in health" workshops with plain language guides so patients can critically assess AI advice.

Protecting patient data

- Regulators should require privacy impact assessments before AI tools are used and whenever data use changes
- Health authorities should mandate audit logs, rapid breach alerts, and patient rights to view, correct, or delete data
- Professional bodies should set clear protocols for documenting AI assisted decisions

Further research is needed to assess generative AI's long term effect on patient outcomes and clinician-patient relationships. Without robust accountability mechanisms and transparent reporting of errors, AI risks deepening mistrust and safeguarding failures, even under the strongest regulatory regimes. Embedding patient and clinician feedback loops will be crucial to catch unintended harms and maintain the integrity of the patient-provider partnership. Understanding the dynamics of use will be essential for designing AI systems that enhance patient autonomy and improve trust in AI assisted care. Without a comprehensive approach that includes regulation, education, data protection, and research, AI in healthcare risks exacerbating inequalities rather than addressing them.

Key messages

- Generative AI can transform patients' access to expertise, but it must also enable them to retain meaningful choice and the capacity to act on that information
- Trust and benefits vary across user groups, and digital divides risk exacerbating existing inequities in access and uptake
- Peer reviewed, patient centred studies are needed to assess the effect of generative AI in clinical and patient facing settings
- Safeguards such as certification of accuracy, transparent data governance, and support for AI literacy are essential to protect and empower patients

Contributors and sources: CB's research focuses on the ethical and social implications of digital health technologies. ML has expertise in primary care applications of digital tools. SR works on digital self-care and lives with Parkinson's disease. DFN has a research focus on digital health policy. RL has interests in evidence based medicine and patient centred care. This article draws on peer reviewed literature, policy documents, and the authors' professional and lived experiences. All authors contributed to writing and reviewing the articles and CB is the guarantor.

AI use: Generative AI (mostly Claude 4 Opus, but also ChatGPT) was used as an initial soundboard to expand on the original series ideas and article structure; to craft a cohesive narrative avoiding textual duplication and providing clarity and conciseness, and to help craft and refine use cases and examples.

Public and patient involvement: SR's experience of living with Parkinson's disease has directly shaped the framing and interpretation of the issues discussed. CB also has two siblings with myotonic dystrophy, and her experiences inform her research and perspective.

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that SR is a member of The BMJ's patient panel.

Provenance and peer review: Commissioned; externally peer reviewed,

1 Pompliano A. My friend went to the doctor recently. X, 21 Mar 2025. https://x.com/APompliano/status/1903205422435103207

2 Holohan M. A boy saw 17 doctors over 3 years for chronic pain. ChatGPT found the diagnosis. Today. 12 Sep 2023. https://www.today.com/health/mom-chatgpt-diagnosis-pain-rcna101843

3 Fraile Navarro D, Lewis M, Charlotte B, etal. Generative AI and the changing dynamics of clinical consultations. *BMJ* 2025;391:e085325.

4 Brodeur PG, Buckley TA, Kanjee Z, et al. Superhuman performance of a large language model on the reasoning tasks of a physician. arXiv 2024:arXiv:2412.10849.v3. [Preprint.]

5 Ayers JW, Poliak A, Dredze M, etal. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med* 2023;183:-96. doi: 10.1001/jamainternmed.2023.1838 pmid: 37115527

6 Kharko A, McMillan B, Hagström J, etal. Generative artificial intelligence writing open notes: A mixed methods assessment of the functionality of GPT 3.5 and GPT 4.0. *Digit Health* 2024;10:20552076241291384. doi: 10.1177/20552076241291384 pmid: 39493632

7 Siddals S, Torous J, Coxon A. "It happened to be the perfect thing": experiences of generative AI chatbots for mental health. *Npj Ment Health Res* 2024;3:. doi: 10.1038/s44184-024-00097-4 pmid: 39465310

8 Presiado M, Montero A, Lopez L, et al. *KFF health misinformation tracking poll: artificial intelligence and health information.* KFF, 2024. https://www.kff.org/public-opinion/kff-health-misinformation-tracking-poll-artificial-intelligence-and-health-information/

9 Menezes MCS, Hoffmann AF, Tan ALM, etalConsortium for Clinical Characterization of COVID-19 by Electronic Health Records. The potential of Generative Pre-trained Transformer 4 (GPT-4) to analyse medical notes in three different languages: a retrospective model-evaluation study. *Lancet Digit Health* 2025;7:-43. doi: 10.1016/S2589-7500(24)00246-2 pmid: 39722251

10 Jin Y, Chandra M, Verma G, et al. Better to ask in English: cross-lingual evaluation of large language models for healthcare queries. *Proceedings of the ACM Web Conference* 2024:2627-38.

11 Morley J, Hine E, Roberts H, etal. Global health in the age of AI: charting a course for ethical implementation and societal benefit. *Minds Mach* 2025;35:doi: 10.1007/s11023-025-09730-3.

12 Decker H, Trang K, Ramirez J, etal. Large language model based chatbot vs surgeon-generated informed consent documentation for common procedures. *JAMA Netw Open* 2023;6:e2336997. doi: 10.1001/jamanetworkopen.2023.36997 pmid: 37812419

13 Inzlicht M, Cameron CD, D'Cruz J, Bloom P. In praise of empathic AI. *Trends Cogn Sci* 2024;28:-91. doi: 10.1016/j.tics.2023.12.003 pmid: 38160068

14 Morley J, Hine E, Roberts H, etal. Global health in the age of AI: charting a course for ethical implementation and societal benefit. *Minds Mach* 2025;35:-35. doi: 10.1007/s11023-025-09730-3.

15 Shemtob L, Majeed A, Beaney T. Regulation of AI scribes in clinical practice. *BMJ* 2025;389:. doi: 10.1136/bmj.r1248 pmid: 40541424

16 Herrman J. ChatGPT wasn't supposed to kiss your ass this hard. Intelligencer. 1 May 2025. https://nymag.com/intelligencer/article/chatgpt-chatbot-ai-sycophancy.html

17 Goddard J. *Hallucinations in ChatGPT: a cautionary tale for biomedical researchers* . *Am J Med* 2023;136:-60. doi: 10.1016/j.amjmed.2023.06.012

18 Eichenberger A, Thielke S, Van Buskirk A. A case of bromism influenced by use of artificial intelligence. *Ann Intern Med Clin Cases* 2025;4:e241260. doi: 10.7326/aimcc.2024.1260

19 Zack T, Lehman E, Suzgun M, etal. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. *Lancet Digit Health* 2024;6:-22. doi: 10.1016/S2589-7500(23)00225-X pmid: 38123252

20 Fox S. Rare disease in the US. 2025. SSRS. 2025. https://ssrs.com/insights/rare-disease-in-the-us-2025/

21 Tyson A, Pasquini G, Spencer A, et al. 60% of Americans would be uncomfortable with provider relying on AI in their own health care. Pew Research Centre, 2023. https://www.pewresearch.org/science/2023/02/22/60-of-americans-would-be-uncomfortable-with-provider-relying-on-ai-in-their-own-health-care/

22 Royal College of Radiologists. The future of AI in healthcare: public perceptions of AI in Radiology. 2025. https://www.rcr.ac.uk/media/fbfn1mxr/the-future-of-ai-in-healthcare-report-appendix.pdf

23 Spain to impose massive fines for not labelling AI-generated content. *Reuters* 11 Mar 2025. https://www.reuters.com/technology/artificial-intelligence/spain-impose-massive-fines-not-labelling-ai-generated-content-2025-03-11/

24 Narayan A, Bhushan I, Schulman K. India's evolving digital health strategy. *NPJ Digit Med* 2024;7:. doi: 10.1038/s41746-024-01279-2 pmid: 39414924