npj | digital medicine

Comment

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-025-01656-5

# Lessons from Henrietta Lacks inform a transparency framework to catalyze generative artificial intelligence in medicine

Tej D. Azad, Anmol Warman, Deven McGraw & Suchi Saria



The integration of generative artificial intelligence (AI) tools into healthcare poses significant challenges concerning data privacy and governance. Drawing on the historical vignette of Henrietta Lacks, this perspective examines the implications of using generative AI in clinical settings. We discuss current health data governance practices and their potential limitations in the generative AI era. We propose a framework of proactive transparency to preserve patient autonomy without limiting technologic progress.

The emergence of large language models (LLMs) has stimulated imaginations regarding the role of generative artificial intelligence (AI) in healthcare. The scale of LLMs endows them with versatility and generalizability, core aspects of their appeal and potential use. The immense quantity of training and tuning data required to achieve state-of-the-art performance in clinical settings has raised concerns about data governance, privacy, and ownership. While many of these concerns extend broadly to clinical research utilizing de-identified patient data or registry databases, generative AI and LLMs in particular, introduce unique challenges due to their scale, complexity, and the dynamic nature of their training processes, including emergent, unanticipated behaviors that require ongoing monitoring rather than a one-time ethics review. Further, LLMs rely on massive data inputs from multiple institutions and data brokers, making provenance increasingly opaque, and their training on such vast corpora introduces unique re-identification risks. Through inspection of historical precedent, namely the example of Henrietta Lacks, we propose a practical path forward to realize the promise of generative AI in healthcare.

In 1951, a Black woman named Henrietta Lacks was treated for cervical cancer. Her cells were cultured and disseminated without consent, giving rise to the immortalized HeLa cell line, which was foundational for biomedical research and led to commercial success. This case not only spurred crucial policy reforms aimed at rectifying racial injustice and exploitation in biomedical research, such as the revised Common Rule, enhanced informed consent protocols for biospecimen use, and increased community engagement guidelines, but also raised enduring questions about the secondary use of data[1]. However, while these reforms improved consent practices and promoted equitable treatment of biospecimens, they were designed for static biological samples and traditional research paradigms, falling short of addressing the novel challenges presented by generative AI and LLMs where data from thousands or millions of patients are aggregated, often invisibly, to train large models. The speed, scale, and opacity of generative AI systems create vulnerabilities that traditional policies, and even policies directed at other modern machine learning approaches, were never designed to mitigate, amplifying ethical and governance concerns in ways that extend beyond the original scope of the HeLa case.

The ethical dilemmas embedded in Lacks' story foreshadow potential pitfalls of deploying generative AI in medicine. If legitimate concerns regarding data provenance are not proactively addressed, we run the risk of ignoring the lessons from Henrietta Lacks and failing to realize potential benefits of generative AI in healthcare. Here, we overview the current state of health data governance and posit that active transparency and thoughtful regulation provide a path forward that engages patients and clinicians without hindering technologic progress.

## Current state of health data governance

A common belief is that patients should simply own their own data. Although intuitive, data ownership is an inadequate approach to preserve patient rights while also allowing for the processes, AI-powered or not, required to improve patient care[2]. Strict data ownership models can stifle innovation and limit the sharing of critical insights that drive improved outcomes. Ownership approaches pose massive technical and logistical challenges to operationalize, can inhibit the interoperability, and limit aggregation of data needed for powering LLMs. Existing privacy and human subjects research rules in the United States largely reject such an approach. These rules require entities who possess identifiable data, or conduct research, to be held accountable for adopting safeguards to protect patient privacy and assure ethical conduct of research, regardless of who may "own" the data. Current privacy laws, albeit imperfect, provide an initial framework for addressing large-scale data uses and disclosures in healthcare. The Health Insurance Portability and Accountability Act (HIPAA) delineates protocols for treatment, population health management, and research utilization of identifiable data, alongside criteria for when data are considered de-identified and therefore not subject to regulation. Similarly, the Common Rule prescribes conditions for employing identifiable data in research. And yet, these statutes fall short of confronting two primary challenges illuminated by the Lacks controversy—opacity and ethical scrutiny deficits.

While HIPAA attempts to embed transparency mechanisms through the Notice of Privacy Practices, these merely outline legal data usage possibilities without requiring entities using the data to disclose actual practices. Similarly, while the Common Rule mandates ethics reviews for research, whether biospecimens or digital data, these focus on interventional research risks without fully addressing the ethics of commercialization, particularly when using de-identified data. Greater priority must be given to refining transparency practices in a manner that provides clarity towards the legality and ethics of data use, ensuring that both legal and moral commitments to patients are upheld.

## A call for practical transparency

A path forward for AI in healthcare requires transparency and proactive communication with patients. As the HeLa case vividly demonstrates, the absence of transparency can lead to long-lasting harm and disenfranchisement. Transparency can be operationalized by extending to patients the practice of disclosing comprehensive information about secondary use of their clinical data, akin to the legally mandated disclosures for research participants[3]. Although this approach is broadly relevant to any clinical research utilizing de-identified data, it is particularly critical for generative AI applications, where the scale and multiplicity of secondary data uses may be obscured by algorithmic complexity. This heightened transparency should encompass scenarios in which a patient's data, even after de-identification, may be used by generative AI tools to improve the patient's own care, drive quality improvement, and inform population health decision-making. Such practice could foster parity in information sharing and contributes to a foundation of transparency and trust. This approach is crucial in balancing the dual objectives of protecting patient privacy and enabling technology adoption. While there is a risk that transparency about AI in healthcare engenders pushback from patients, the loss of trust accompanying any perception of deceit by patients is likely far more detrimental, both at the health system level and at the broader technologic adoption level. Preliminary studies suggest that patients are still developing their own perspectives on AI in healthcare, further reinforcing the need for proactive communication by trusted healthcare practioners[4].

Distinct from primarily research-focused institutional review boards (IRBs) and more practical than direct patient-level data ownership, independent data ethics review boards may offer a step towards addressing ethical considerations in the use of health data[5]. These boards would assess projects involving data collection, sharing, or sale—weighing benefits and risks, societal impact, and privacy policies. Importantly, care must be taken to ensure these boards retain true independence, rather than serving as an institutional mechanism to limit external data access or rubberstamp internal data requests. Independence can be achieved by mandating a governance structure that includes external, non-affiliated members with veto power, strict conflict of interest policies, and the mandatory public disclosure of review outcomes. Additionally, these boards should be structured to incorporate external oversight through independent community advisory panels, cross-institutional legal experts, and periodic external audits. Such a multi-layered approach establishes a robust counterbalance to internal biases, ensuring decisions are continuously scrutinized beyond institutional confines, thereby maintaining the board's integrity and bolstering public trust in data practices - an imperative underscored by the historical lessons of the HeLa case, which illustrated the particularly lasting impact of opaque practices on marginalized communities.

Increasing transparency is a distinct approach to consent-based recommendations. Transparency involves clearly communicating to patients how their data may be used and ensuring accountability in data practices, whereas consent involves obtaining explicit permission for specific data uses. When considering the role of patient consent for data use by generative AI systems, we posit that HIPAA provides an initial framework. Under HIPAA, consent is required when identifiable health data is used for research, but not when used for quality improvement or care delivery. Consent can also be waived by an Institutional Review or Privacy Board weighing privacy risks against the importance of the research. Data fed into generative AI tools can be governed similarly, focusing on regulation of the use case, not the underlying technology. Importantly, consent is not required by HIPAA for the use, or re-use, of de-identified data—this should not be considered differently simply because AI is being used. In applications with clear patient benefit, not requiring consent for de-identified data can remove an unnecessary, dauting barrier to innovation. However, in scenarios where benefits are ambiguous, such as frontier research, it may prove necessary to consider a structured consent process. Tiered consent models that empower patients to choose varying levels of data use could be considered, complemented by dynamic consent frameworks that allow ongoing adjustments as research evolves. To best align with patient priorities and ensure that ethical practices reflect diverse stakeholder values, active community engagement is necessary to guiding the construction of consent tiers.

The primacy of transparency and the protection of patients in the implementation of AI in healthcare is not merely a procedural necessity; it is integral to the cultivation of trust—especially among marginalized communities. Members of historically underserved communities, who may have well-founded skepticism about a blurred line between clinical care and medical research, deserve clear communication about how their data may be stored, used, and commercialized.

Addressing these concerns involves proactively informing patients about the use of AI, including AI's role in clinician decision-making and the continuous learning nature of these technologies. This commitment to transparency should be tailored to the specific state of AI implementation. In settings where AI functions in a fully autonomous capacity, it is essential to clearly delineate decision-making boundaries and establish robust accountability measures. Alternatively, in purely assistive models, emphasizing the collaborative role between clinicians and AI is crucial. In both cases, articulating the degree of human oversight, potential risks, and limitations of the system can help maintain patient trust while ensuring that ethical considerations are contextually grounded. While similar transparency measures could benefit any data-intensive research, the unique characteristics of generative AI demand a more sophisticated ethical apparatus - one that is agile enough to adapt to evolving data practices while steadfastly upholding the values of patient autonomy and trust. If done well, the integration of generative AI into clinical medicine takes us closer to a future where access to appropriately de-identified health data can be democratized for public good. If done without a commitment to transparency, there is a risk of repeating the harms done to Henrietta Lacks, at scale, and setting back technological advances in healthcare.

Despite the benefits of a practical transparency framework, several limitations warrant consideration. The complexity and cost of implementing external oversight mechanisms, such as independent review boards with multi-stakeholder representation, may be prohibitive for smaller institutions or emerging research initiatives. Additionally, maintaining dynamic consent processes and regular audits requires substantial infrastructure, ongoing training, and technical expertise, which may strain existing resources. Mitigation strategies include establishing centralized oversight bodies that serve multiple institutions, leveraging digital platforms to streamline consent management, and securing dedicated funding streams to support continuous monitoring and evaluation. Even with robust governance and dynamic consent, fully understanding how these models internalize and use data remains difficult, necessitating that oversight strategies continuously incorporate ongoing research into model interpretability. These measures can help ensure that the ethical ambitions of transparency do not compromise operational feasibility or innovation.

Striking the delicate balance between innovation and ethical considerations is a moral imperative. We should commit to prioritizing transparency to safeguard patients and build a foundation of trust that ensures the equitable and ethical improvement of medical care.

## Data availability

No datasets were generated or analysed during the current study.

**Tej D. Azad**[1,2] ✉, **Anmol Warman**[3], **Deven McGraw**[4] **& Suchi Saria**[5,6]
[1]Department of Neurosurgery, Johns Hopkins Hospital, Baltimore, MD, USA. [2]Hopkins Business of Health Initiative, Baltimore, MD, USA. [3]Johns Hopkins University School of Medicine, Baltimore, MD, USA. [4]Citizen, San Francisco, CA, USA. [5]Johns Hopkins University, Baltimore, MD, USA. [6]Bayesian, New York, NY, USA. ✉e-mail: tazad1@jhmi.edu

Received: 3 December 2024; Accepted: 21 April 2025;
Published online: 14 May 2025

## References

1. Beskow, L. M. Lessons from HeLa cells: the ethics and policy of biospecimens. *Annu. Rev. Genom. Hum. Genet.* **17**, 395–417 (2016).
2. Evans, B. J. Much ado about data ownership. *Harv. J. Law Technol.* **25**, 69–130 (2011).
3. Spector-Bagdady, K. Hospitals should act now to notify patients about research use of their data and biospecimens. *Nat. Med.* **26**, 306–308 (2020).
4. Aggarwal, R., Farag, S., Martin, G., Ashrafian, H. & Darzi, A. Patient perceptions on data sharing and applying artificial intelligence to health care data: cross-sectional survey. *J. Med. Internet Res.* **23**, e26162 (2021).
5. Parasidis, E., Pike, E. & McGraw, D. A Belmont report for health data. *N. Engl. J. Med.* **380**, 1493–1495 (2019).

## Author contributions

T.D.A., A.W., D.M., S.S. participated in the writing and revision of the manuscript. All authors have read and approved the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Tej D. Azad.