Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AI in Healthcare: Ethical Considerations in Patient Data Management

Dr. Rohith Vangalla

Optum Services, Inc., USA

ARTICLE INFO

Received: 27 Sept 2025

Revised: 29 Oct 2025

Accepted: 09 Nov 2025

ABSTRACT

The introduction of artificial intelligence into healthcare provision is both revolutionary and raises significant ethical issues that should be closely considered. The issue of patient data management seems to be one of the critical concerns when the problem of privacy vulnerability, breach of security, and unauthorized access jeopardizes the most basic rights of the people who receive medical services. Another notable issue is algorithmic bias, since the training sets that are not demographically diverse will result in systems that promote historical healthcare disparities and provide poor care to the underrepresented groups. Many artificial intelligence systems are opaque and therefore lack the transparency to support informed consent and add complexity in clinical decision-making. Black box algorithms, which do not understand how they make their decisions, cause a loss of trust between healthcare providers and patients and lead to safety concerns when unsolicited recommendations are followed. To develop sustainable artificial intelligence healthcare systems, sustainable models of governance to involve the patients as active stakeholders instead of passive sources of data are needed. Health care institutions should be accountable, show transparency in reporting, honest opt-out, and provide tangible proof of the ability to enhance the quality and outcome of care with the help of algorithmic tools. Educational programs, which clarify the artificial intelligence abilities and constraints, will enable patients to be actively involved in decision-making, resulting in their treatment and data use. It is necessary to balance technological novelty with ethics in the development of reliable systems that would benefit every population in a fair manner. The way forward requires a continuous interaction between technologists, clinicians, policymakers, ethicists, and patient advocates to set the standards that can safeguard individual rights and, at the same time, facilitate positive innovation. Technical sophistication itself will not be used to measurably gauge success, but rather how far artificial intelligence systems support the principles of human dignity, fairness, and autonomy in healthcare delivery.

Keywords: Artificial Intelligence In Healthcare, Patient Data Privacy, Algorithmic Bias, Healthcare Transparency, Collaborative Governance

1. Introduction

Healthcare stands at a peculiar junction where machines are beginning to think alongside doctors, nurses, and medical staff. Fortune Business Insights has tracked how hospitals and clinics worldwide are pouring money into artificial intelligence tools, betting that these technologies will solve problems that have plagued medicine for decades [1]. From diagnostic imaging software that spots tumors to predictive systems that forecast patient deterioration, AI applications are spreading through healthcare facilities at a breakneck pace. Accenture's investigation into this phenomenon reveals something interesting: these tools aren't just fancy gadgets—they're addressing real headaches like inconsistent treatment approaches, mountains of paperwork, and missed diagnoses that cost lives and money [2]. But here's where things get complicated. All advances in AI healthcare appear to carry some ethical baggage, particularly over patient data management. In medical records, there are some

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

of the most personal facts of the lives of individuals, including embarrassing symptoms and genetic predispositions, and mental health issues. As algorithms begin to process millions of these records, the questions that can emerge are those that were absent ten years ago. How much access is too much? Who decides what's fair? Will the patients be assured that their information will not be abused? It is not a philosophical discussion about this or that abstract idea. It is a practical issue of real-life people seeking medical treatment. Having this balance is critical to the future of medicine.

2. Patient Privacy and Data Security Challenges

Medical data has become a peculiar kind of gold rush in the digital age, except instead of prospectors with pickaxes, there are hackers with keyboards trying to break into hospital databases. The IBM Security Cost of a Data Breach Report paints a troubling picture: when healthcare organizations get hacked, the financial hit dwarfs what other industries face, and that's before counting the damage to reputation and patient confidence [3]. Why does healthcare make such an attractive target? Medical records fetch premium prices on the black market—a complete health history can sell for ten times what a stolen credit card number brings. One of the issues can be traced to the development of the healthcare IT systems. Hospitals did not create their networks with security in mind, but rather have pieced them together over the decades, having different security standards, all of which are tied to dozens of third-party vendors. Factor in the fact that emergencies occur, at times, doctors require patient information promptly, and security measures that slow things down are bypassed or compromised. Now throw AI into this mess. Training a decent medical algorithm requires feeding it thousands or millions of patient records. More data generally means better performance, creating pressure to aggregate information from multiple sources. But aggregation means concentration, and concentration means a single breach exposes more people. Abouelmehdi and colleagues explored how blockchain technology might help, suggesting that decentralizing health data storage could prevent the kind of massive breaches that make headlines [4]. Blockchain leaves permanent documents of users and files viewed by them in the past, and at what time, unauthorized snooping can be easily identified. The privacy regulations could be automatically implemented through smart contracts, and do not need to be monitored by a person at all times. A few hospitals have been trying these methods, but not very many are using them. In the meantime, there is a new direction in methods such as federated learning in which algorithms are trained in more than one location without data being centralized. Differential privacy introduces mathematical noise to data, which serves to safeguard personal identities without undermining the general trends. The synthetic data is generated to provide artificial patient records that replicate or follow the real patterns to be used in training. Each approach has tradeoffs. Such laws as HIPAA provide minimum expectations, yet most professionals claim that those laws were designed in a pre-AI world and do not sufficiently address the existing risks.

Challenge Category	Description	Impact
Data Breach Vulnerability	Healthcare records command premium prices on the black market; fragmented IT infrastructure creates multiple entry points	Financial losses exceed other industries; long-term reputation damage and patient trust erosion
Training Data Requirements	AI algorithms require access to massive patient datasets for effective performance	Creates tension between data utility needs and confidentiality obligations
Third-Party Access	Multiple vendors and interconnected systems increase exposure points	Expanded attack surface for potential breaches and unauthorized access

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Emergency Access Protocols	Urgent medical situations sometimes bypass security measures	Weakened security protocols create ongoing vulnerabilities
Regulatory Frameworks	Existing laws like HIPAA predate modern AI applications	Compliance meets minimum standards but may inadequately address emerging risks

Table 1: Privacy and Security Challenges in AI Healthcare Systems [3, 4]

3. The Critical Issue of Data Bias and Algorithmic Fairness

The most dangerous biases are sometimes the ones that no one should have set up. Obermeyer and colleagues stumbled onto a shocking discovery while examining a healthcare algorithm used by hospitals and insurance companies across America: the system was essentially telling Black patients to get sicker before deserving the same level of care that white patients received [5]. Here's the twist—nobody programmed racism into the algorithm. Instead, developers used healthcare spending as a shortcut for measuring health needs, figuring that sicker people cost more to treat. Sounds logical, right? It overlooked an important fact, however: Black patients have traditionally received lower-quality medical care and have created lower healthcare expenses not due to their medical condition, but due to barriers to receiving medical care in the system. The algorithm has then been trained on this biased history and reproduces it on a large scale, which involves millions of people. This was not a little glitch in a little program; this was a massively relied upon system making resource allocation decisions, which dealt with life and death. Bias creeps into medical AI through countless pathways. Most training datasets skew heavily toward certain demographics because most medical research has historically focused on specific groups. When dermatology AI trains mostly on images of light skin, it struggles to diagnose conditions in darker skin. When cardiac models train predominantly on male patients, they miss heart attack symptoms that present differently in women. Sjoding and colleagues found bias hiding in an even more basic medical device: the pulse oximeter, which estimates blood oxygen levels [6]. These tools inaccurately estimate high oxygen levels in dark-skinned patients, as these algorithms were not designed or tested with sufficient diversity. The consequences? Late detection of a threatening oxygen level may result in more adverse consequences. What makes algorithmic bias particularly insidious is the scale. A biased human doctor might harm dozens or hundreds of patients over a career. A biased algorithm can make thousands of flawed decisions daily across multiple healthcare systems simultaneously. Fixing this requires more than good intentions. Healthcare organizations need to actively seek out diverse datasets, even when that's harder and more expensive than using convenient samples. Algorithms need ongoing monitoring across different demographic groups, not just overall accuracy metrics. When bias surfaces—and it will surface—systems need rapid response protocols. Some researchers are developing standardized fairness tests for medical AI, similar to how drugs undergo clinical trials before approval. The problem is that even fairness is a disputable area, and there are various definitions of fairness, which are contradictory with one another.

Bias Source	Mechanism	Consequence
Historical Healthcare Inequities	Algorithms trained on past spending patterns perpetuate access disparities	Underrepresented groups require higher illness severity for equivalent care recommendations
Demographic Dataset Gaps	Training data overrepresents certain populations while excluding others	Reduced diagnostic accuracy and inappropriate treatment suggestions for minority groups

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Device Algorithm Limitations	Medical devices developed without diverse testing populations	Systematic measurement errors affecting patients with darker skin pigmentation
Proxy Variable Selection	Using neutral-seeming metrics that encode historical bias	Seemingly objective algorithms institutionalize discrimination at scale
Validation Testing Gaps	Performance was evaluated on overall accuracy rather than demographic subgroups	Bias remains undetected until real-world deployment affects vulnerable populations

Table 2: Sources and Manifestations of Algorithmic Bias in Healthcare [5, 6]

4. Transparency and Informed Consent in AI-Driven Healthcare

Black boxes work fine for music streaming recommendations; they're terrifying in medical contexts. According to studies released in Digital Health, there is an essential conflict of interest: the AI-based medical decision-making machines are frequently too complicated to be comprehended by anyone, including their developers [7]. Physicians are in a very awkward situation when an algorithm proposes a treatment and the patient asks why, then the truthful one may say: The computer says so and I really do not know how it came to that conclusion. Several issues arise out of this transparency. Patients are entitled to make rational judgments regarding their future and bodies. Trust is the key to doctor-patient relationships that demand transparency. Patients can inquire of a physician as to his or her rationale, the basis, and alternatives in case he or she prescribes surgery or medicine. What happens when the reasoning is buried inside algorithmic calculations involving thousands of weighted variables? Some healthcare providers simply don't mention AI involvement, figuring it's easier to avoid complicated explanations. That approach sidesteps short-term awkwardness but corrodes trust when patients eventually discover algorithms influenced their care without their knowledge. Other providers mention AI but provide vague reassurances rather than meaningful information. Neither approach respects patient autonomy. Challen and colleagues pointed out that opacity creates safety risks beyond ethics [8]. When the doctors are unable to comprehend the reason why an algorithm suggests a certain thing, they are unable to properly assess whether the suggested recommendation would be reasonable given the specifics of a given patient. Blind trust in algorithmic outputs leads to medical errors—situations where unusual patient factors that would have prompted a human to question a recommendation go unnoticed because the algorithm spoke and everyone assumed it must be right. Some types of AI are more explainable than others. Simple decision trees can be traced step by step. Deep neural networks remain largely mysterious. Scientists are coming up with methods such as attention mechanism that indicates which variables had the greatest impact on a prediction, or counterfactual explanations on how the change in certain inputs would impact outputs. These instruments are beneficial but not the solution. Informed consent has been regarded as a signature, once, on a form before a procedure. AI alters this calculus since algorithms are not fixed, the use of data is wider, as well as applications appear following preliminary approval. Significant consent in an AI scenario entails discussions that are sustained and made transparent, as well as real chances to decline with no consequences.

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Transparency Issue	Problem	Potential Solution
Black Box Algorithms	Complex systems resist explanation even by creators	Explainable AI techniques, including attention mechanisms and counterfactual reasoning
Limited Patient Understanding	Technical complexity prevents meaningful informed consent	Ongoing dialogue and accessible explanations in plain language
Clinician Knowledge Gaps	Providers cannot explain AI reasoning to patients or evaluate appropriateness	Training programs and decision support tools that surface algorithmic logic
Safety Risks	Inability to question recommendations leads to unchallenged errors	Transparency requirements that enable critical evaluation by medical professionals
Evolving Applications	New uses emerge after initial consent	Continuous consent processes rather than one-time authorization

Table 3: Transparency Barriers and Solutions in AI Healthcare [7, 8]

5. Building Trust Through Patient Engagement and Collaborative Governance

Trust is a strange thing—hard to build, easy to break, and essential for healthcare. Richardson and colleagues examined what actually influences whether patients feel comfortable sharing health data and found that trust doesn't come from technical features alone [9]. Individuals desire to be informed that organizations will be responsible, that data exchange will have purposes that they believe are valuable to them, and that there is real control in the way information is used. In many cases, patients in healthcare institutions have been viewed as sources of data to be exploited instead of stakeholders who need to be engaged. Here, the consent process is made a formality in order to get things through. Inquiries are discouraged as employees are on the move. The issue of privacy is brushed off with vague promises of security. This is a way of creating compliance and not trust. Other progressive healthcare systems are trying out new models. Instead of feedback on emerging issues, patient advisory boards consider proposed AI projects and data use policies and provide feedback before the project is implemented. Such boards have a wide range of representation based on different ages, ethnicity, socioeconomic status, and health status, since homogeneous groups lack critical views. The community forums involve the audience in a discussion of ethical limits on the use of health data. What applications feel acceptable? Where should lines be drawn? Interactive educational programs help demystify AI, explaining both capabilities and limitations in plain language. Collaborative governance brings multiple voices to the table: clinicians using AI tools daily, technologists building them, ethicists identifying moral considerations, patient advocates representing affected communities, and researchers studying impacts. Davenport and Kalakota observed that technical sophistication means little if human factors like trust, usability, and workflow alignment get neglected [10]. The brightest algorithms are collecting dust where they are not trusted by doctors or patients refuse to use them. Trust is not achieved by making promises, but whenever one is to build trust, he/she need to be accountable in a tangible way. AI system performance, failures, and limitations should be recorded well through regular transparency reports. There should have been clear channels of complaints with visible evidence that the concerns are addressed. The alternative to opt-out should be that real patients who refuse the use of some data should not be subjected to inconspicuous punishment or low-quality care. Above all, healthcare organizations are expected to have a visible impact of AI adoption. Does care quality improve? Do outcomes get better? Is it a reduction in costs or

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

an increased access? Trust is built because individuals view that the risks they take yield rewards that they appreciate. Education programs must not be hysterical and fear-mongering. Artificial intelligence has actual possibilities and actual constraints. Patients should receive genuine evaluations of the two. With knowledge of what algorithms can and cannot accomplish, people are better placed to engage in decision-making in a meaningful manner, as it relates to the integration of AI into their care.

Engagement Strategy	Implementation	Benefit
Patient Advisory Boards	Include diverse representation in AI development decisions	Incorporates multiple perspectives before implementation rather than after problems emerge
Transparent Reporting	Document AI system performance, including failures and limitations	Demonstrates accountability and honest communication about capabilities
Genuine Opt-Out Options	Allow data usage decline without care quality penalties	Respects autonomy and provides meaningful control over information
Educational Initiatives	Explain AI capabilities and limitations in accessible terms	Empowers informed decision-making and reduces fear or unrealistic expectations
Tangible Benefit Demonstration	Show measurable improvements in care quality and outcomes	Builds confidence that accepted risks produce valued rewards

Table 4: Elements of Trust-Building Through Patient Engagement [9, 10]

Conclusion

The medical field has reached a pivotal point in the development of artificial intelligence, as technical functions have significantly surpassed moral standards, which drives the necessity to develop holistic solutions, in which patient rights and fair care are central to innovation. The protection of privacy requires more than regulatory compliance to include well-built security architectures, access controls based on blockchains, minimization of data, such as federated learning and differential privacy, which allow algorithmic development without sacrificing the confidentiality of individuals. The response to algorithmic bias should actively include working to create multiple training datasets, deploy ongoing monitoring of demographic subgroups, develop quick response strategies to instances of disparities, and define standardized measures of fairness to provide a consistent level of equity evaluation across applications and institutions. Transparency issues require a shift towards explainable artificial intelligence systems capable of explaining a line of reasoning in ways accessible to regular people, and rethink approaches to consent that consider patients as lifelong partners and not signatories to a form they hardly comprehend. To foster trust, it is necessary to build collaborative governance frameworks that involve clinicians, technologists, ethicists, patient advocates, and community representatives to define the implementation of artificial intelligence by adopting a common set of values, instead of being strictly technical. Healthcare institutions should be accountable by showing their performance reporting, avenues of complaints, authentic opt-out procedures, and visible indications that algorithmic technologies do bring the promised changes in care quality and patient outcomes. Educational programs need to be truthful about the abilities and limitations of artificial intelligence technologies, with no hype to create false expectations and fearmongering to block positive innovation. Artificial intelligence healthcare systems will not become trustworthy, fair, and truly beneficial to all patient groups despite demographic traits to become the ultimate measure of success. Technology must be used to benefit humanity and not vice versa, and upholding that ideal will define whether the potent tools will create more health equity or will be used to reinforce the existing inequalities and cause new types of harm targeting vulnerable populations in the most perilous ways.

Journal of Information Systems Engineering and Management
2025, 10(62s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

References

[1] Fortune Business Insights, "AI in Healthcare Market Size, Share & Industry Analysis, By Platform (Solutions and Services), By Application (Robot-Assisted Surgery, Virtual Nursing Assistant, Administrative Workflow Assistance, Clinical Trials, Diagnostics, and Others), By End-user (Hospitals & Clinics, Pharmaceutical & Biotechnology Companies, Contract Research Organization (CRO), and Others), and Regional Forecasts, 2025-2032," 2025. [Online]. Available: https://www.fortunebusinessinsights.com/industry-reports/artificial-intelligence-in-healthcare-market-100534

[2] Accenture Consulting, "Artificial Intelligence: Healthcare's New Nervous System". [Online]. Available: https://www.accenture.com/content/dam/accenture/final/a-com-migration/manual/r3/pdf/pdf-49/Accenture-health-artificial-intelligence-j.pdf

[3] IBM Security, "Cost of a Data Breach Report 2022". [Online]. Available: https://www.key4biz.it/wp-content/uploads/2022/07/Cost-of-a-Data-Breach-Full-Report-2022.pdf

[4] Anton Hasselgren et al., "Blockchain in healthcare and health sciences—A scoping review," International Journal of Medical Informatics, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S138650561930526X

[5] Ziad Obermeyer et al., "Dissecting racial bias in an algorithm used to manage the health of populations," Science, 2019. [Online]. Available: https://www.science.org/doi/10.1126/science.aax2342

[6] Michael W. Sjoding et al., "Racial Bias in Pulse Oximetry Measurement," The New England Journal of Medicine, 2020. [Online]. Available: https://www.nejm.org/doi/full/10.1056/NEJMc2029240

[7] Kerstin Noelle Vokinger et al., "Digital health and the COVID-19 epidemic: an assessment framework for apps from an epidemiological and legal perspective," Swiss Med Wkly, 2020. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/32418194/

[8] Robert Challen et al., "Artificial intelligence, bias and clinical safety," BMJ Quality & Safety. [Online]. Available: https://qualitysafety.bmj.com/content/28/3/231

[9] Kathrin Cresswell et al., "Health Care Robotics: Qualitative Exploration of Key Challenges and Future Directions," J Med Internet Res., 2018. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29973336/

[10] Thomas Davenport and Ravi Kalakota, "The potential for artificial intelligence in healthcare," Future Healthc J., 2019. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC6616181/