Journal of Applied Social Sciences and Humanities

https://journal.krrishpub.com/index.php/jassh/index

RESEARCH ARTICLE

Automated Indifference Artificial Intelligence and Human Rights in Healthcare

Yichuan Wang [a,*]

[a] *Associate Professor, Management School, University of Sheffield, Sheffield, United Kingdom ,S102TN , United Kingdom.*

Abstract

The integration of Artificial Intelligence (AI) into global healthcare systems represents a seismic shift in medical epistemology, yet it resurrects Gilbert Ryle's philosophical concept of the "ghost in the machine" in a troubling new form. No longer a critique of dualism, the "ghost" has mutated into the unexamined biases and opaque logic structures codified within deep learning networks. This paper investigates the tension between "algorithmic determinism" and fundamental human rights, specifically focusing on non-discrimination and informed consent. Utilizing a secondary data analysis methodology, the study synthesizes critical findings from 2021 through early 2025, including the landmark Estate of Lokken v. United Health Group class-action lawsuit and recent data on diagnostic error rates. The research identifies a "responsibility gap" where moral agency is ceded to probabilistic outputs, exemplified by the nH Predict algorithm's 90% error rate in denying post-acute care. The findings suggest that without a radical restructuring of ethical governance—moving beyond voluntary principles to enforceable human rights impact assessments—healthcare risks descending into a state of automated indifference. Key themes include the erosion of the fiduciary relationship, the "black box" barrier to consent, and the systemic erasure of marginalized demographics in predictive modeling.

Keywords: Algorithmic bias, Medical ethics, Human rights, Black box medicine, Informed consent, AI Governance.

Article information:
Published by: **Krrish Scientific Publications Pvt. Ltd.**
DOI: https://doi.org/10.71426/jassh.v1.i1.pp15-19
Received: *30 Nov. 2025* | Revised: *26 Dec. 2025* | Accepted: *30 Dec. 2025* | Published: *31 Dec. 2025*

1. Introduction

The machinery of modern medicine is haunted. In 1949, philosopher Gilbert Ryle introduced the phrase "the ghost in the machine" to dismantle the idea that a non-physical mind sits inside a mechanical body, pulling the levers. It was a critique of separation. Today, as we stand on the precipice of a healthcare revolution driven by AI, Ryle's metaphor has returned with a vengeance. The "machine" is now the vast infrastructure of digital health—electronic records, diagnostic scanners, and insurance databases. The "ghost" is the invisible, often unintelligible logic of the algorithm itself, a spectral force that influences life-and-death decisions through mechanisms that are often opaque even to their creators.

The urgency of this issue is driven by the sheer scale of the technological takeover. The global market for AI in healthcare is projected to surpass significant milestones by 2030, fueled by the promise of hyper-efficiency and the democratization of expert diagnostics. However, this rapid deployment has revealed deep fissures in the ethical bedrock of medicine. The World Health Organization (WHO) [10] and legal scholars have issued repeated warnings that the unchecked expansion of AI threatens to exacerbate existing health inequities, creating a system where the "digital divide" becomes a determinant of survival. The core of the crisis lies in the tension between the clinical mandate to "do no harm" and the technological imperative to optimize for efficiency. The AI models learn from historical data, they inevitably ingest the prejudices of the past. If the history of healthcare delivery is marred by racism or classism, the AI trained on that history becomes an agent of those biases, effectively automating discrimination under the guise of mathematical neutrality. Furthermore, the "black box" nature of deep learning models poses a direct challenge to the human right of informed consent. How can a physician explain the risks of a treatment if the rationale is locked inside a proprietary neural network? This paper seeks to navigate these dark waters, asking the fundamental question: In the age of the algorithm, who owns the decision to heal?

*Corresponding author
Email address: yichuan.wang@sheffield.ac.uk (Y Wang).

2. Literature review

The scholarly discourse surrounding AI ethics in healthcare has shifted sharply between 2021 and 2025, moving from optimistic manifestos to critical empiricism and legal scrutiny. This review synthesizes recent literature to map the contours of this evolving debate. The resurgence of Ryle's "ghost" serves as a potent framework for recent commentators. Hashmi [1] utilizes the metaphor to discuss the integration of AI in medical education, arguing that the "ghost" represents the elusive human element—empathy and communication—that risks being exorcised by an over-reliance on data-driven pedagogy. This aligns with Milossi et al. [2], who explore "algorithmic determinism," or the idea that human agency is being eroded by predictive analytics. They suggest a shift in the locus of control where the physician's authority is slowly usurped by the machine, creating an epistemological crisis. If the machine operates on a logic distinct from human reasoning, the human operator is reduced to a passive bystander.

A substantial volume of research focuses on the entrenchment of bias. Unlike early studies that viewed bias as a glitch, recent work by Anderson and Visweswaran [3] frames it as a structural inevitability of current data practices. Their scoping review on "algorithmic individual fairness" highlights how models trained on unrepresentative datasets fail to account for individual variations, effectively erasing minority experiences. This is further complicated by the "predictive" nature of these tools. Narayanan and Kapoor [4], in their critique of "AI Snake Oil," argue that we have conflated the impressive capabilities of generative AI with the dubious accuracy of predictive AI. They warn that using algorithms to predict social outcomes—like whether a patient will be "non-compliant"—is fundamentally flawed and scientifically suspect.

The legal landscape is also grappling with the "black box" problem. The class-action lawsuit Estate of Gene B. Lokken v. United Health Group has become a central case study in this domain. Birrane et al. [5] detail how the lawsuit challenges the use of the nH Predict algorithm, which allegedly used rigid statistical averages to deny care to elderly patients, overriding the judgment of treating physicians. This marks a turning point where algorithmic ethics moves from academic debate to federal litigation. Finally, Schiff [6] raises a nuanced concern about the "dehumanization" of clinical notes, fearing that AI-driven documentation will strip the medical record of the narrative nuance that often holds the key to accurate diagnosis. The consensus in the literature is that we are facing a "responsibility gap," where the complexity of technology is used to evade accountability for medical harm. Similar views are echoed by other authors [11] –[15]. The following objectives were set for the study:

1. To investigate the collision between the theoretical promise of AI in healthcare and the practical reality of its deployment.
2. To quantify the impact of algorithmic bias on the right to non-discrimination and to evaluate the "black box" problem as a barrier to informed consent.

3. Research methodology

This study employs a secondary data analysis methodology, which is particularly investigating the rapidly evolving intersection of bioethics and law. Given the proprietary nature of many medical algorithms, direct primary data collection is often impossible for external researchers. Secondary analysis allows us to triangulate data from diverse, high-authority sources to construct a holistic picture of the ethical ecosystem.

Data was aggregated from sources published between January 2021 and November 2025. We utilized legal filings from the Lokken litigation to analyze error rates in insurance algorithms, peer-reviewed meta-analyses for diagnostic accuracy statistics, and recent governance reports from the WHO. The analytical framework is guided by a Human Rights-Based Approach to Data (HRBAD), which asks not just what the data says, but who is harmed by it. We extracted quantitative statistics regarding error rates and adoption barriers and synthesized legal arguments to construct a narrative about human agency. The primary limitation is the reliance on reported data, as the internal code of these "black boxes" remains a trade secret.

4. Analysis

The analysis of the collected data reveals the mechanisms through which the "ghost in the machine" operates, specifically through the trade-off between efficiency and safety. Recent studies have sought to benchmark AI performance against human specialists. Data synthesized from 2024 radiology reports (referencing Radiology and related benchmarking studies) illustrates this comparison.

The data in Table 1 exposes a complex reality. While the domain-specific AI model was statistically "preferred" in ranking, its error rate in localization (14.8%) was actually slightly higher than that of human radiologists (13.9%). More concerning is the performance of general-purpose models like GPT-4 Vision, which had a staggering 32.7% error rate. This illustrates the danger of deploying non-specialized "predictive" tools in high-stakes environments. The "ghost" here is the illusion of competence; the AI produces a confident-sounding report that may be factually incorrect, a phenomenon Schiff [6] warns could contaminate the medical record with "hallucinations."

Table 1: Comparative diagnostic performance (AI vs. Human radiologists)

Metric	Domain-Specific AI Model	Human Radiologists	GPT-4 Vision (General AI)
Error Rate (Localization)	14.8%	13.9%	32.7%
Clinical Significance of Errors	56.5%	65.8%	73.3%
Preference Ranking (1st Place)	60.0%	–	–
Preference Ranking (2nd Place)	–	54.7%	–

Table 2: Algorithmic determination in insurance (The *nH Predict* case) [8]

Metric	**Statistic**	**Context**
Algorithm Name	nH Predict	Used by UnitedHealthcare / NaviHealth
Alleged Error Rate	90%	Percentage of denials overturned on appeal
Appeal Rate	0.2%	Percentage of patients who actually appeal
Outcome	Denial of post-acute care	Overrides physician recommendation

Table 2 presents the most damning evidence of algorithmic human rights violations. A 90% error rate—meaning the algorithm was wrong nine out of ten times it was challenged—is statistically indistinguishable from random chance, yet it was used to systematically deny care to the elderly. This is "algorithmic determinism" weaponized. The algorithm did not predict recovery; it predicted the corporate desire to limit spending. The fact that only 0.2% of patients appeal reveals the mechanism of harm: the system relies on patient exhaustion. The "ghost" is the corporate entity hiding behind the machine, using the complexity of the code to evade responsibility for what is effectively a breach of contract.

Data from 2025 surveys highlights why, despite these risks, the human firewall is crumbling.

Table 3: Workforce barriers to ethical AI Adoption

Barrier Category	**Prevalence**	**Impact on Human Rights**
Lack of AI Knowledge	High (universal concern)	Inability to obtain valid informed consent
Automation Bias	Moderate to high	Erosion of the right to health due to deference to algorithmic error
Liability Fear	High	Defensive medicine and the emergence of a “responsibility gap”

Table 3 illustrates the fragility of the human element. The fear of liability and a lack of technical knowledge creates "automation bias," where clinicians defer to the machine's judgment even when they suspect it is wrong. This surrender of agency severs the fiduciary bond between doctor and patient.

The synthesis of legal filings and clinical data leads to definitive findings. First, Institutionalized Discrimination is a feature, not a bug. The high error rates in generalist models and insurance algorithms confirm that these systems are often "functionally blind" to the nuance of individual patient needs, validating the Bias Hypothesis. The 90% overturn rate in the nH Predict case proves that these tools can systematically disadvantage vulnerable populations (the elderly) to optimize financial metrics.

Second, we observe the Collapse of Informed Consent. The "black box" problem described in [7] is not merely technical but ethical. If a physician cannot understand why an AI recommended a specific treatment (or denial of care), they cannot convey the risks to the patient. Consequently, the patient cannot give valid consent. The patient is asked to trust a system that is legally protected as a trade secret, violating their right to information and bodily autonomy.

5. Discussion

These findings force a confrontation with the philosophical underpinnings of our healthcare system. Ryle's "ghost" was a critique of the idea that a mind could control a body from the outside. We have inverted this; we have built a digital body—the AI infrastructure—and trapped a ghost inside it. This ghost is a "shadow self" comprised of our collective biases and profit motives.

The deployment of tools like nH Predict reveals a shift from Deontological duty (care for the individual) to Utilitarian efficiency (care for the system). The algorithm optimizes for the average, but human rights exist to protect the outlier. When we automate decision-making, we automate indifference. The "Responsibility Gap" identified in the literature is the death knell of trust. If an AI makes a mistake, the developer blames the data, the hospital blames the vendor, and the insurer blames the algorithm. The patient is left with no one to sue and no one to heal them. We are creating a system where no one is to blame because "the computer said so."

Findings align with similar studies by Gore and Olawade [9]. They resonate in general findings from other research [16] – [20]. AI decision-making is not a wholly new concept, but it is one that has blossomed in prevalence in recent years, especially with its unexpected spike in popularity and use cases. The ethical implications of AI raise concerns about more than just mistakes or blunders as it progressively influences decisions, whether they are made by humans or are entirely autonomous. Where should artificial intelligence be applied? What is the boundary? AI's potential to make judgments for us touches on the fundamental problems of human autonomy, fairness, bias, and our trust as a species in an ever-more automated world that will continue to affect how we live our lives.

Since AI's conception, the ethics issue has been raging. There isn't a single right way to apply it, nor is there a catch-all answer. Fundamentally, it is a technology that needs to be handled carefully and pragmatically. However, this does not answer the critical questions surrounding its ability to make judgments, especially decisions that might resonate across our society: who bears accountability when AI makes mistakes? How can we train these systems on large datasets while maintaining individual privacy? Accountability frameworks and equitable implementation techniques must be carefully considered in order to strike a delicate balance between expanding AI capabilities and preserving human autonomy. Perhaps most of all, it requires a comprehensive knowledge of the ethical issues that come with its usage.

A basic and regrettable paradox at the core of AI ethics is that as AI systems grow more sophisticated and powerful, their decision-making processes may become less apparent to human control. This is the core principle underpinning what is known as the 'black box' problem. This problem is especially common with deep learning AI models that make use of intricate neural networks, where data is processed through numerous layers of connected nodes and inputs are given "tokens" that organize data in a hierarchical manner.

The issue here is that it is frequently impossible to ascertain how an AI arrives at its findings. The input and the output are known to us, but everything in between is

unknown. Numerous ethical and technological issues arise from the inability to discern how an AI model arrives at any particular conclusion. Technically, this means that if there is an issue with an AI's output, it is tough to tell where things are going wrong and what to change. Complex AI is a different animal than normal software, where problems are easier to find and resolve with a patch.

The people it makes decisions for may suffer greatly as a result of algorithmic prejudice. There have been instances of algorithmic discrimination, such as the denial of healthcare to people of color or the preference of men over women for employment. Even if inadvertent, these prejudices reinforce already-existing disparities and, if uncontrolled and completely trusted, might have had—and have had—seriously negative consequences for individuals.

AI models may be biased due to a variety of factors, including human error from things like faulty testing procedures, underrepresentation in the development teams that produce AI models, and historical datasets that reflect current cultural prejudices. As AI decision-making becomes more prevalent, it is imperative to maintain fairness and integrity.

Although algorithmic bias has been a problem in the past and will surely continue to occur occasionally, it has not gone undetected. This can be a difficult problem to resolve because of the black box issue. However, AI models can eventually become more bias-aware with ongoing monitoring, increased human oversight, and bias detection technologies.

Many would contend that this is "the big one," and they might not be entirely incorrect. The biggest concerns, especially for businesses wishing to adopt AI, is how AI models, acquire, use, and learn from data. Due to concerns about data abuse, several major players, including Apple, Verizon, and iHeartRadio, have prohibited the use of models like ChatGPT. Samsung, in particular, limited the usage of the chatbot after discovering that employees had entered crucial code into it.

One crucial ethical boundary is the possibility of unintentional corporate trade secret exposure. AI systems might mistakenly leak sensitive information and create major legal and competitive hazards for businesses. Sophisticated data governance strategies that safeguard both individual privacy and business intellectual property are essential for organizations looking to integrate AI components, particularly generative AI.

Generative AI models are trained on both input data and data scraped from the internet at a broader societal level, which means that they can—and frequently do—memorize all of the material that is provided to them, regardless of sensitivity. When paired with things like biometric data, expressions, and other personally identifiable information (such as financial records and credit scores), individuals may find themselves unwittingly exposed, with intimate personal characteristics transformed into computational data points without meaningful consent. Opaque algorithmic evaluations that reduce human complexity to numerical scores may significantly limit an individual's economic potential.

As previously stated, human supervision and ongoing monitoring are now essential components of controlling AI. Making sure that human-in-the-loop procedures are in place is essential to more efficient, risk-averse, and advantageous AI decision-making. These frameworks protect human judgment at important decision points while exploiting AI's processing capabilities. AI has a lot to offer businesses, but risk mitigation calls for human intervention.

Where AI is used, a cost/benefit ratio study is necessary. When a decision-making function has little to no wider impacts, full automation is more appropriate. Businesses that plan to use AI should develop a set of best practices for their staff that strike a balance between the effectiveness of AI and corporate responsibility. One such example is oversight thresholds.

For further protection, a strong AI governance system can be added to human control. A methodical strategy that incorporates ethical design principles, interdisciplinary oversight, ongoing monitoring, and established accountability systems will help reduce risks at the governance level from the outset. For particularly risk-averse organizations, building multi-tiered review processes, embedding ethical considerations directly into any and all AI-based designs, and establishing adaptive governance models, corporations may ensure that AI systems remain fundamentally subject to human judgment. For additional reassurance, the framework should provide visible audit trails and require explicit chains of accountability.

The way forward is one that demands excitement but cautious. AI is an incredibly effective instrument with almost infinite possible uses. It is simple to become enthralled with this technology and use it without performing the necessary research. To guarantee that risks are reduced, organizations deploying AI systems should make investments in strong oversight procedures with precise deployment criteria. As previously mentioned, AI is an excellent tool, but it is still a tool. A hammer may only be as harmful or safe as the person using it permits.

6. Conclusion

The "ghost in the machine" is real, dangerous, and currently unregulated. The integration of AI poses a clear danger to the human rights of patients globally. We conclude that bias is baked into the infrastructure, opacity nullifies consent, and efficiency is being prioritized over ethics.

The implications are severe. Policymakers must abolish the "trade secret" defense for medical algorithms; if an algorithm makes a medical decision, its logic must be auditable. Clinicians must be trained to recognize "automation bias" and empowered to override the machine. Future research must move from diagnosis to treatment, developing "Human Rights Impact Assessments" that are mandatory before any AI is deployed in a clinical setting. The machine is here to stay, but we must rewrite the code to respect the dignity of the human spirit.

Declarations and ethical statements

Conflict of interest: The author declares that there is no conflict of interest.
Funding statement: The author declares that no specific funding was received for this research.

Artificial Intelligence usage statement: During the preparation of this work, the author used ChatGPT to assist with grammatical corrections. After using this tool, the author reviewed and edited the content as needed and take full responsibility for the content of the published article.

Availability of data and materials: Data used from secondary sources have been cited in the article.

CRediT authorship contribution statement

Yichuan Wang: Conceptualization, Investigation & Writing - Original draft.

Publisher's note

Krrish Scientific Publications Pvt. Ltd. and the ***Journal of Applied Social Sciences and Humanities*** remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Hashmi AM. Ghost in the Machine: Artificial Intelligence in Medical Education. *Annals of King Edward Medical University* 2025 Jun 30;31(Spl2):97-98. Available from: https://doi.org/10.21649/akemu.v31iSpl2.6163

[2] Milossi M, Alexandropoulou-Egyptiadou E, Psannis KE. AI ethics: algorithmic determinism or self-determination? The GPDR approach. *Ieee Access*2021 Apr 12;9:58455-66. Available from: https://doi.org/10.1109/ACCESS.2021.3072782

[3] Anderson JW, Visweswaran S. Algorithmic individual fairness and healthcare: a scoping review. *JAMIA Open*. Volume 8, Issue 1, February 2025, ooae149,Available from: https://doi.org/10.1093/jamiaopen/ooae149

[4] Narayanan A, Kapoor S. AI Snake Oil : What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference *Princeton University Press* 2025. Available from: https://www.torrossa.com/en/resources/an/6055496

[5] [Online Available]: Birrane K, Tobey D, Kopans D, Cloud W. Lawsuit over AI usage by Medicare Advantage plans allowed to proceed. *DLA Piper*(2025, February 24). Available from: https://www.dlapiper.com/en/insights/publications/ai-outlook/2025/lawsuit-over-ai-usage-by-medicare-advantage-plans-allowed-to-proceed

[6] Schiff GD. AI-Driven Clinical Documentation—Driving Out the Chitchat?. *New England Journal of Medicine* 2025 May 15;392(19):1877-9. Available from: https://doi.org/10.1056/NEJMp2416064

[7] Chau M, Rahman MG, Debnath T. From black box to clarity: Strategies for effective AI informed consent in healthcare. *Artificial Intelligence in Medicine*. 2025 May 24:103169. Available from: https://doi.org/10.1016/j.artmed.2025.103169

[8] [Online Available]: Estate of Gene B. Lokken et al. v. UnitedHealth Group, Inc. et al. Case No. 0:23-cv-03514 (D. Minn. 2023). Available from: https://litigationtracker.law.georgetown.edu/wp-content/uploads/2023/11/Estate-of-Gene-B.-Lokken-et-al_20231114_COMPLAINT.pdf

[9] Gore MN, Olawade DB. Harnessing AI for public health: India's roadmap. *Frontiers in Public Health*. 2024 Sep 27;12:1417568. Available from: https://doi.org/10.3389/fpubh.2024.1417568.

[10] [Online Available]: World Health Organization. Ethics and governance of artificial intelligence for health: large multi-modal models. WHO guidance. World Health Organization; 2024 Jan 18. Available from: https://www.who.int/publications/i/item/9789240084759.

[11] Formosa P, Rogers W, Griep Y, Bankins S, Richards D. Medical AI and human dignity: Contrasting perceptions of human and artificially intelligent (AI) decision making in diagnostic and medical resource allocation contexts. *Computers in Human Behavior*. 2022 Aug 1;133:107296. Available from: https://doi.org/10.1016/j.chb.2022.107296

[12] [Online Available]: Correia PM, Pedro RL, Videira S. Artificial Intelligence in Healthcare: Balancing Innovation, Ethics, and Human Rights Protection.*Journal of Digital Technologies and Law*. 2025;3(1):143-180. Available from: https://doi.org/10.21202/jdtl.2025.7

[13] Parry MW, Markowitz JS, Nordberg CM, Patel A, Bronson WH, DelSole EM. Patient perspectives on artificial intelligence in healthcare decision making: a multi-center comparative study. *Indian Journal of Orthopaedics*. 2023 May;57(5):653-65. Available from: https://doi.org/10.1007/s43465-023-00845-2

[14] van Leersum CM, Maathuis C. Human centred explainable AI decision-making in healthcare.*Science Direct* Journal of Responsible Technology. 2025 Mar 1;21:100108. Available from: https://doi.org/10.1016/j.jrt.2025.100108

[15] Aizenberg E, Van Den Hoven J. Designing for human rights in AI. *Big Data & Society. 2020* Aug;7(2):2053951720949566. Available from: https://doi.org/10.1177/2053951720949566

[16] Molbæk-Steensig H, Scheinin M. Human Rights and Artificial Intelligence in Healthcare-Related Settings: A Grammar of Human Rights Approach.*Brill.com* European Journal of Health Law. 2025 May 13;32(2):139-64. Available from: https://brill.com/view/journals/ejhl/32/2/article-p139_2.xml

[17] Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-assisted decision-making in healthcare: the application of an ethics framework for big data in health and research. *Asian Bioethics Review*. 2019 Sep;11(3):299-314. Available from: https://doi.org/10.1007/s41649-019-00096-0

[18] [Online Available]: Hoxhaj O, Halilaj B, Harizi A. Ethical implications and human rights violations in the age of artificial intelligence.*Balkan Social Science Review*. 2023 Dec 25;22(22):153-71. Available from: https://www.ceeol.com/search/article-detail?id=1207107

[19] Nouis SC, Uren V, Jariwala S. Evaluating accountability, transparency, and bias in AI-assisted healthcare decision-making: a qualitative study of healthcare professionals' perspectives in the UK. *BMC Medical Ethics*. 2025 Jul 8;26(1):89. Available from: https://doi.org/10.1186/s12910-025-01243-z

[20] Elgin CY, Elgin C. Ethical implications of AI-driven clinical decision support systems on healthcare resource allocation: a qualitative study of healthcare professionals' perspectives.*BMC Medical Ethics*. 2024 Dec 21;25(1):148. Available from: https://doi.org/10.1186/s12910-024-01151-8