IBOM MEDICAL JOURNAL
Vol.18 No.3 | July - September 2025 | Pages 440-450
www.ibommedicaljournal.org

Ethical and Legal Dimensions of AI Diagnosis in Medicine

Emmanuel Iniobong Archibong[1], Rukaiya Oyiza Raji[2], Doris Ifeoma Ibe[3], Abigail Emmanuel Udoh[4], Moses Chisunka[5]

[1]Department of Philosophy, University of Uyo, Uyo, Nigeria.
[2]Department of Public Health, University of Portharcourt, Nigeria.
[3]Department of Public Health, Abia State University, Uturu Nigeria.
[4]Department of Social Work and Social Welfare Administration, Ignatuis Ajuru University of Education, Porharcourt Nigeria.
[5]Department of Pharmacy Practice and Public health, School of Pharmacy and Health Sciences, United States International University - Africa, Nairobi, Kenya

Abstract

The recourse to Artificial Intelligence (AI) in medical diagnosis has ambivalent or paradoxical dimensions. It is believed to have transformed healthcare resulting to improved accuracy and efficiency on one hand, while on the other, it raises concerns regarding ethical and legal dimensions. Because of the sensitive issues of the study, a qualitative method of critical/conceptual analysis of literature was employed to understand the intricate dimensions in the use of AI in matters that bothers on the sacredness of life in the face of wrong diagnosis. A fundamental research question that guided the study was: "what are the ethical and legal implications of AI-driven failed diagnosis"? The interdisciplinary approach of the study aimed at examining more closely, the impact of AI diagnosis on medical care alongside ethical and legal liability. Findings revealed threshold of errors, lack of accountability and legal backing as some outcomes in the use of AI for diagnosis. The discussion indicated the need for clear guidelines and regulations governing AI use in medicine from an ethical and legal perspective. The study's significance is in its contribution to the overall understanding of AI-use in healthcare delivery and the guarding against faulty diagnosis. Recommendations from the study include establishing standardized AI protocols which will ensure transparency in AI decision-making in medicine as well as training for healthcare practitioners who use the system.

Keywords: Artificial Intelligence, Medical Diagnosis, Ethics, Legal Backing, Healthcare, Patient Care

Introduction

Artificial Intelligence (AI) is no longer a futuristic concept in healthcare as it has become an embedded element in clinical decision-making systems. In radiology, pathology, dermatology, and cardiology, AI-enabled diagnostic tools are being deployed with the promise of precision, speed, and efficiency. These technologies are supposedly trained to detect anomalies, flag patterns, and provide probabilistic outcomes with a level of consistency that is believed can outperform human interpretation in specific contexts. However, the same systems that can improve diagnostic accuracy also introduce a plethora of ethical dilemmas and legal uncertainties particularly around transparency, patient autonomy, and professional liability.

What makes AI particularly problematic is its operational obscurity as many of the most advanced algorithms function as "black boxes" conclusions through complex machine learning processes that

Corresponding Author:

Emmanuel Iniobong Archibong, Ph.D.
Department of Philosophy, University of Uyo, Uyo, Akwa Ibom State, Nigeria.

emmaiarchibong@uniuyo.edu.ng

DOI: 10.61386/imj.v18i3.702

are not easily interpretable, even by their developers[1]. The clinical value of AI tools is frequently judged based on their predictive accuracy rather than their ability for explanation. As a result, healthcare providers are often left using diagnostic tools whose internal logic they do not fully understand. While AI may enhance diagnostic consistency, its use can simultaneously undermine informed consent and the physician's moral agency.

These concerns have prompted an array of scholarly responses. In underscoring the paradoxical nature of AI diagnostics, it is stated that "While AI technologies promise a revolution in healthcare, they risk becoming ethically unsound if used without rigorous oversight and transparency"[2]. When diagnostic errors occur in AI-assisted decisions, it becomes unclear who is responsible whether the algorithm's developer, the healthcare institution, or the individual clinician. Legal scholars have noted the insufficiency of existing tort backgrounds to accommodate AI-related negligence claims, given that such assume a human actor who can be judged against a reasonable standard of care[3]. Moreover, patients are rarely informed when AI systems are involved in their diagnosis.

Accordingly,[4] most clinical interfaces do not distinguish between human-led and machine-generated diagnostic insights, creating a form of silent substitution that compromises informed consent. This introduces not just ethical concerns, but also serious legal implications. The legal notion of consent hinges on patient knowledge and voluntary agreement to procedures. When AI is silently integrated, patients may unknowingly accept diagnostic outcomes that they would otherwise question or reject if made aware of the technology's involvement. This scenario is further complicated by issues of accountability. In traditional malpractice litigation, courts evaluate whether a physician acted competently given the standard of care. But in AI-assisted cases, what constitutes reasonable care becomes murky. If a physician follows AI advice that turns out to be incorrect, are they negligent? Or would deviating from AI advice have been riskier?

In many jurisdictions, there remains a legal void in addressing AI-specific liabilities. The U.S. Food and Drug Administration (FDA) and European Medicines Agency have initiated preliminary frameworks, particularly in regard to AI as medical devices. However, these regulations focus primarily on safety and efficacy during initial approval, not post-deployment accountability in clinical use[5]. Furthermore, current contexts do not adequately address adaptive AI systems that evolve through continuous learning, meaning their diagnostic logic can shift in ways that are not predictable or testable after release.

The emergence of AI introduces epistemological shifts since traditional diagnosis relies on clinical reasoning, experiential judgment, and patient narratives. AI, in contrast, relies on pattern recognition and correlation across large datasets, often with minimal regard for individual context. This form of diagnosis may be accurate statistically but deficient from a human standpoint. A physician might understand a patient's story, socio-economic context, or emotional needs but AI cannot. As the integration of AI accelerates, these concerns are not diminishing but are intensifying. Legal scholar has cautioned that "regulatory lag is endemic in fast-evolving technological fields. In healthcare, this lag can have life-and-death consequences"[6]. This challenge is what this study will be analysing in order to have a foothold of AI induced diagnosis and its implications for patients and health workers who are both at the centre of it all.

Research Questions

The following research questions will guide the study's logical analysis trajectory viz:

i. To what extent are healthcare practitioners aware of the ethical implications of using AI in clinical diagnosis?
ii. What are the perceived legal concerns among health practitioners regarding AI-assisted diagnosis?
iii. How do patients perceive the use of AI in making medical decisions that affect their care?
iv. Does the implementation of AI diagnostic systems improve the efficiency of healthcare delivery without compromising ethical standards?
v. What is the relationship between healthcare practitioners' trust in AI systems and their willingness to adopt them in clinical practice?

Methodology

This study adopts a qualitative research

methodology, guided by a critical-conceptual analytical approach. This approach is particularly well-suited to exploring the ethical and legal dimensions of Artificial Intelligence (AI) in clinical diagnosis that are normative, interpretive, and situated within socio-legal and philosophical contexts, rather than empirically quantifiable. The rationale for this is not to measure phenomena but to interpret the conceptual architecture that underpins evolving norms around medical accountability, informed consent, professional duty, and technological responsibility. Consequently, it explained that, "the intersection between medical AI, law, and ethics defies traditional empirical frameworks because it is rooted in divergent value systems and regulatory vacuums"[7].

The decision to adopt a qualitative conceptual method is therefore both philosophical and functional. Rather than drawing from statistical trends or numerical models, the research critically analyses academic literature, medical ethics doctrines, legal texts, and international AI governance frameworks to unpack the layers of implication around AI-assisted diagnosis. This method also resonates with prior work in technology ethics. For instance, it has been argued that "conceptual analysis allows researchers to confront the moral assumptions embedded in technology adoption assumptions that often escape scrutiny in data-driven studies"[8]. Their research on mental health AI emphasized the risk of ethical drift in algorithmic implementation where foundational concepts such as "care," "autonomy," or "risk" are operationalized without ethical interrogation. Similarly, this study examines the way such foundational ethical and legal terms are challenged or reshaped when AI is silently integrated into diagnostic environments. In line with methodological transparency, this study aspires to ethical and legal clarity which is a conceptual and normative lens through which healthcare stakeholders (policy-makers, clinicians, patients) can anticipate and address the layered risks of AI diagnosis. Through conceptual and critical engagement with ethical theory, legal doctrine, and interdisciplinary scholarship, the study aims to elucidate the conceptual gaps, normative tensions, and actionable insights needed to shape responsible and human-centred AI deployment in clinical settings.

Conceptual Clarifications

Artificial Intelligence (AI)

Artificial Intelligence in healthcare refers to systems that perform tasks normally requiring human intelligence, such as learning, reasoning, and pattern recognition. Specifically, in diagnostic medicine, AI includes machine learning models, artificial neural networks, and natural language processing systems that interpret imaging, pathology reports, or patient records to suggest or automate diagnosis. The World Health Organization defines AI as "the application of machine learning algorithms and other cognitive technologies to simulate human intelligence processes, particularly for clinical decision support, diagnostics, and automation in healthcare delivery"[9]. These systems are designed to process vast datasets beyond human cognitive capacity, identifying correlations and trends in symptoms, medical images, or histories to generate diagnostic outputs.

In practice, AI algorithms (including machine learning and deep learning models) process large, multimodal datasets such as clinical records, medical images, and bio-signals to recognize patterns and make predictions. However, within clinical settings, AI's definition must also account for its non-neutrality. It has been stressed that, AI is not a passive tool but a "normatively embedded socio-technical system whose design and deployment reflect specific human decisions, values, and limitations"[10]. For instance, an AI diagnostic tool that prioritizes speed over explanation introduces an ethical compromise. Its use shapes the decision environment and influences how clinicians perceive their diagnostic authority. AI is here treated not merely as a computational device but as a clinical actor whose recommendations are often given epistemic and practical authority in decision-making. This status complicates questions of responsibility, especially when the system is not clear and probabilistic rather than deterministic.

Medical Diagnosis

Medical diagnosis traditionally refers to the clinical reasoning process through which a healthcare professional identifies a disease or condition based on the evaluation of a patient's symptoms, signs, test results, and history. It is a blend of art, science, and

judgment requiring both empirical observation and intuitive synthesis. Diagnosis involves "the interpretation of symptoms and signs using clinical knowledge to reach a conclusion about the most likely cause of illness"[11]. AI however, shifts this process by automating data interpretation and proposing differential diagnoses based on statistical patterns, rather than experiential judgment.

This mechanization introduces both promise and peril. AI can outperform clinicians in repetitive or image-heavy tasks such as skin lesion classification or CT scan anomaly detection. AI technologies are increasingly integrated into this diagnostic workflow. AI systems can automatically analyze radiology images (X rays, CTs, MRIs, ultrasounds, etc.) to detect disease markers that might be missed by the human eye. It has been argued that, this process lacks the "contextual integration of social, emotional, and psychosomatic cues that are often crucial in diagnosis"[12]. AI systems operate by correlation, not causation, which can lead to technically accurate but clinically inappropriate recommendations. Thus, in this study, medical diagnosis is treated as a clinically situated act that may be enhanced but also potentially disrupted by AI systems. When AI is deployed without clear epistemic boundaries or transparency, it blurs the lines between supportive analysis and decisive determination.

Ethics

In every culture, there is a metaphysics in so far as people engage with "ultimate questions of reality. By ultimate questions of reality is meant, general puzzling or thought provoking without a clear answer from experimental science"[13]. Ethics provides ultimate questions about the rightness and wrongness of an action by an agent. In healthcare, ethics refers to the normative framework guiding right and wrong conduct in medical practice, rooted in values such as autonomy, beneficence, non-maleficence, and justice. Bioethics has historically served as the guiding lens for assessing medical decisions, especially in life-altering contexts such as end-of-life care, reproductive health, or experimental treatments. In the context of AI, ethics expands to include issues like transparency, accountability, algorithmic bias, and patient consent in data-driven decisions. It has been observed that, "AI challenges traditional ethical norms because it introduces non-human decision agents that lack moral accountability yet exert significant clinical influence"[14]. Informed consent, for instance, is ethically compromised when patients are unaware that an AI system generated the diagnostic recommendation upon which treatment is based.

The World Health Organization outlines six core ethical principles for AI in health: autonomy, transparency, accountability, inclusiveness, equity, and sustainability[15]. These principles underpin the evaluation of whether an AI system is ethically fit for deployment. However, ethical breaches often arise not from the failure of principles themselves but from their institutional neglect such as when cost pressures override consent procedures or when clinicians are inadequately trained to interpret AI outputs. In this study, ethics is approached as both a regulatory aspiration and a moral imperative a set of standards that must evolve alongside technological integration, ensuring that innovation does not erode human dignity, care, or professional judgment.

Law

Law refers to the codified systems of rules that regulate medical practice, patient rights, and healthcare accountability. In AI-assisted diagnostics, legal considerations include liability for harm, data protection, regulatory compliance, and the enforceability of informed consent when decisions are machine-influenced. One of the most pressing legal concerns is accountability in diagnostic error. It has been noted that, "legal doctrine is fundamentally anthropocentric as it presupposes a human agent who can be assessed against a standard of reasonableness or duty of care. AI disrupts this structure by diffusing agency across systems, developers, clinicians, and institutions"[16].

Literature Review

Artificial Intelligence (AI) has quickly emerged as a key factor in medical diagnostics, with remarkable potential to transform healthcare through speed, precision, and pattern recognition. However, its integration has exposed ethical and legal pitfalls that demand closer examination. It is held that, "AI introduces a third actor into medical decisions, one that does not bear responsibility in the traditional legal sense"[17]. This observation exposes a legal

vacuum; unlike physicians, AI tools cannot be sued, and manufacturers often protect themselves through disclaimers or distributed accountability models. Consequently, responsibility often reverts to clinicians or institutions even when they rely in good faith on FDA-approved algorithms.

It is argued that the existing legal frameworks "are anthropocentric and ill-equipped to accommodate distributed agency" in decision-making contexts[18]. Courts, typically apply the 'reasonable physician' standard; yet what is deemed reasonable when an AI tool outputs a diagnosis that a physician neither fully understands nor contradicts? The danger here is twofold: clinicians may become over-reliant on algorithms to avoid personal error, while patients are left without clear avenues for redress when AI-driven misdiagnoses occur. This dilemma also has institutional implications with hospitals increasingly bearing secondary liability as the "deployers of intelligent agents,"[19] creating a grey zone in which responsibility diffuses across technical, human, and organizational layers.

Consent procedures in traditional medicine assume the patient is informed about who or what is making decisions affecting their health. AI complicates this norm. just as a "a significant proportion of AI diagnostic systems are deployed without disclosing their role to patients"[20], eroding informed consent from a procedural right into a nominal formality. If a patient consents to a treatment based on a diagnosis they believe comes from a human expert, but which was actually generated by a machine-learning model, is that consent valid? Again, "autonomy is compromised not only when choices are restricted, but also when decision-making contexts are opaque"[21]. This issue is particularly critical in vulnerable populations, such as the elderly or illiterate, who may not question or comprehend the role of AI.

Some scholars argue for the institutionalization of "algorithmic transparency protocols"[22] to ensure patients are notified when AI contributes to diagnosis even though current legal regulations, including GDPR's right to explanation, are unevenly applied across jurisdictions and rarely enforced in real-time clinical settings. As a result, ethical protections lag behind technological realities. Another area of significant concern is the embedded bias within AI diagnostic systems. These systems are only as good as the data they are trained on, and medical data often reflect historical inequalities in access, diagnosis, and treatment. For instance, training datasets that underrepresent Black patients in dermatological images lead to poorer diagnostic performance for skin conditions in non-white populations. It is also noted along this line that "retinal diagnostic algorithms showed significantly lower accuracy for patients with darker eye pigmentation due to skewed training data"[23].

These findings point to a tension between efficiency and equity. While AI can reduce error for majority populations, it may worsen outcomes for minorities. From a legal perspective, this raises the possibility of unequal impact claims, although proving algorithmic discrimination is often difficult due to the opacity of exclusive models. This is called "justice gap of AI diagnostics" which is a structural inequity that is not always visible but often measurable in outcomes[24]. Without a transparent audit, patients or their advocates cannot prove that an algorithm failed them in a legally actionable way.

One of the paradoxes of AI integration is its potential to both augment and erode clinical judgment. While AI may support faster and more comprehensive diagnostic suggestions, it can also shift the epistemic authority away from physicians toward machines. A scholar warns that, "doctors may become 'explainers' of decisions they did not actually make," which fundamentally alters the clinician-patient relationship[25].

It has been argued that when AI becomes the dominant source of diagnostic authority, it undermines the concept of relational care, in which the patient is understood as more than just data points[26]. Law is yet to catch up with this shift as regulatory frameworks do not yet ask whether AI systems should be designed to defer to, rather than dominate, clinical discretion. This is particularly urgent in cases involving probabilistic or low-confidence outputs. A system may flag a likely diagnosis with 72% confidence, but what guidance is offered to clinicians about how to integrate this into their duty of care? The rapid adoption of Artificial Intelligence (AI) in medical diagnosis has been driven by claims of efficiency and accuracy, but beneath these promises lies a complex terrain of risks. These risks are neither incidental nor fully technological as they arise from how AI intersects

with clinical judgment, legal responsibility, and ethical obligations. One of the foremost problems is the opacity of AI systems, especially those built on deep learning. These systems often function as "black boxes," offering conclusions without accessible justifications just as ".... machine learning as a subset of artificial intelligence focuses on developing algorithms that enable computers to learn from data and improve their performance overtime without being explicitly programmed"[27].

In diagnostic settings, this opacity undermines clinical transparency. It has been argued that, "current AI models lack explainability, meaning that physicians often cannot determine why a diagnostic recommendation was made"[28]. This gap challenges the core principle of informed clinical decision-making, where reasoning, not just outcomes, is essential. A system that predicts lung cancer with 95% accuracy but offers no rationale creates a dilemma raising questions whether a physician can responsibly follow it without understanding it? Opacity is not just a technological inconvenience but an ethical deficit. When clinicians defer to systems they cannot audit, and when patients are treated based on opaque logic, the right to explanation is diluted.

Rather than reducing error, AI may reconfigure it, shifting from human misjudgement to overreliance. This behavioural dependency undermines the notion that AI will augment human intelligence and in poorly governed contexts, it supplants critical thinking with passive acceptance. Automation bias has further consequences in liability assessment. If a misdiagnosis arises from clinician deference to a faulty algorithm, is the clinician negligent, or are they following protocol? The inability to trace responsibility across human and machine components renders traditional legal accountability models inadequate. Algorithmic bias, originating from unrepresentative training data is another major challenge. In healthcare, this bias can have fatal consequences, particularly for underserved or minority populations. It has been noted that "AI models trained on predominantly white, male populations routinely underperform for Black patients and women"[29]. This structural bias is not merely a matter of statistical deviation but has become embedded discrimination when deployed at scale. Inaccurate diagnosis of cardiac conditions in women, or late detection of melanoma in dark-skinned patients, are not just clinical failures but ethical violations of equity and justice.

The above point is reinforced by the fact that "bias is not just an artefact of data as it is a reproduction of societal patterns of neglect, now coded into algorithms"[30]. The legal system currently lacks the tools to prosecute this kind of harm, as discriminatory outcomes from AI rarely meet the threshold for direct culpability. As a result, patients harmed by biased algorithms often have no legal remedy unless a clinician acted exceptionally. The technological nature of the error is seen as incidental rather than systemic. Beyond individual patient harm, bias creates disparities in population-level health outcomes. If AI systems misdiagnose more frequently in low-income or rural populations, these groups will suffer increased morbidity due to misclassification and delayed treatment. This effect undermines public health objectives and calls into question whether AI can truly democratize healthcare.

The preference for human judgment is rooted not in irrationality but in the belief that human clinicians can be reasoned with, held accountable, and understood. AI lacks these features as it does not communicate empathy nor can be cross-examined, and rarely offers interpretability. As a result, its clinical authority is both powerful and fragile: it commands trust when hidden and loses it when revealed. AI also raises technical limitations that are often downplayed. Many diagnostic models are trained on ideal datasets that do not reflect real-world complexity. They may fail when confronted with missing data, comorbidities, or atypical presentations. Such inconsistencies make it dangerous to extrapolate lab performance accurately without extensive validation. Again, the rise of AI brings with it data governance risks. These systems require massive datasets to function effectively, raising concerns about patient privacy, consent, and secondary use of data. Patients may therefore consent to a diagnostic procedure without realizing their data will be used to retrain or update AI systems globally, potentially in ways that affect other patients or generate commercial value.

The Sanctity of Human Life

At the core of this study is the sanctity of human life

where the integration of Artificial Intelligence (AI) into medical diagnostics compels a re-examination of its core ethical doctrine. The sanctity of human life is a principle embedded in both secular and religious medical traditions. This principle asserts that human life holds inherent, non-negotiable value, independent of utility, efficiency, or probability. AI, particularly in contexts of diagnosis, challenges this moral anchor by shifting clinical reasoning from person-centred care to data-driven logic. This section critically explores how the deployment of AI in diagnostic medicine confronts, reconfigures, or in some cases, undermines the sanctity and moral inviolability of human life.

The sanctity of life is traditionally grounded in the belief that every human being possesses intrinsic worth. It has been argued that "in Catholic moral theology, the inviolability of life is non-instrumental and cannot be suspended by efficiency or technological promises"[31]. This suggests that the act of diagnosing, especially when misdiagnosis could lead to harmful or fatal outcomes, is more than clinical to being fundamentally moral. A misdiagnosis rendered by an AI system is not simply a computational failure; it becomes a moral injury when it diminishes the patient's right to accurate and empathetic care. What makes this tension urgent is that AI systems operate on probabilistic risk models, which inherently trade off some false positives and negatives to optimize overall accuracy.

It is noted that, "AI does not seek to preserve life for its own sake but to optimize diagnostic predictions based on aggregate performance"[32]. This utilitarian logic is at odds with the traditional ethos of medicine, where the preservation of each individual life is paramount, even if costly or inconvenient. The sanctity of life resists this reduction, asserting instead that each misdiagnosis no matter how statistically rare represents a grave ethical failure when avoidable. Moreover, AI systems often depersonalize the diagnostic encounter, reducing patients to data points. If AI systems replace or overshadow human engagement in the diagnostic moment, the patient may feel processed rather than cared for. This mechanization of medical ethics dilutes the humanistic core of clinical care.

Even in non-lethal diagnoses, such as identifying developmental disorders, AI may inadvertently pathologize or stigmatize. It has been emphasized that "from the Catholic view, human dignity persists even amid imperfection or disability; AI must not become a tool for normative categorization"[33]. When AI renders diagnoses based on neural scans or behavior profiling, the line between clinical assessment and moral judgment blurs. If not ethically constrained, such systems risk undermining the moral worth of those who diverge from algorithmic norms. In addition to philosophical concerns, the religious traditions that shaped bioethics provide further insights. It has been observed that "any application of AI in medicine must proceed from an unwavering respect for the human person not as a computational subject but as a bearer of sacred worth"[34]. This theological view does not oppose innovation but insists that technological advancement be anchored in compassion, not merely precision. AI, they argue, should assist healing, not rationalize detachment.

Many AI systems have been found to favour majority health profiles thereby making invisible the specific risks to minorities and disabled patients. When diagnostic tools are not trained on diverse datasets, they violate the ideals of sanctity by structurally excluding certain lives from accurate attention. Some critics argue that promoting sanctity may impede technological progress. However, the principle does not deny innovation as it demands discernment. It is explained that, "the sanctity of life principle is not anti-technology; it seeks to re-centre empathy and suffering at the heart of machine-enabled healthcare"[35]. AI developers and clinicians ought to know how design choices, training data, and deployment settings either reinforce or erode this empathy. Engineering decisions very well become ethical decisions when lives are involved. The sanctity of life principle holds that even in suffering, life retains moral value and cannot be reduced to predictive analytics.

Positive Laws and the Safeguarding of Human Lives

Another important implication of this study is that, the increased use of Artificial Intelligence (AI) in medical diagnostics creates legal ambiguities regarding responsibility, regulatory enforcement, and patient protection. While ethics provides normative direction, positive law which is a formal statutes, legal doctrines, and codified frameworks grounds the enforceable safeguards for human life in healthcare. The core legal problem AI introduces is

that it complicates the attribution of responsibility and the enforcement of rights when diagnostic decisions lead to harm.

Legal traditions in medical malpractice centre on the duty of care between patient and physician. AI introduces a non-human actor, raising questions about whether existing malpractice laws can accommodate machine-based reasoning. It has been argued that "the Bolam test, which assesses negligence based on peer-accepted standards, becomes inadequate when clinical decisions originate from black-box algorithms rather than peer clinicians"[36]. If the standard is based on what a reasonable doctor would do, it cannot be applied coherently to machine logic that lacks interpretability or peer consensus. This legal disjuncture exposes patients to risk and clinicians to uncertain liability. Efforts to clarify civil liability for AI-induced harm have been made in some jurisdictions. It has been suggested that "civil law must recognize AI as a risk-amplifying actor, akin to a hazardous product, thus justifying strict liability doctrines for developers and institutions"[37]. If a diagnostic AI system errs and a patient suffers harm, liability would attach to the party that placed the system in clinical use, regardless of fault. This is a shift from traditional malpractice to product liability logic, reflecting AI's dual status as a tool and an agent.

International legal instruments are also evolving and the World Health Organization (WHO), in its 2021 guidance on AI in health, emphasizes that "states have an obligation under international human rights law to ensure that health-related AI technologies do not result in harm and that accountability mechanisms are in place to redress harm when it occurs"[38]. This places positive duties on governments, not just developers or clinicians. National regulatory bodies must not only approve AI systems but also monitor their real-world performance, audit their outcomes, and establish remedies for wrongs. A critical area of concern is data protection, which guard the legality of AI training and deployment. For instance, The European General Data Protection Regulation (GDPR) enshrines the right to explanation (Article[22], yet few healthcare AI systems offer this feature. Without meaningful transparency, patients are denied full disclosure, breaching both ethical informed consent and legal standards for valid decision-making.

Positive law must also address interjurisdictional uncertainty, particularly in telemedicine and cross-border diagnostics. A radiologist in France may rely on an AI system developed in the United States, deployed via a cloud server based in India, diagnosing a patient in Nigeria. In such cases, whose law governs? It has been noted that "as AI diagnostic services become cloud-based and international, legal accountability is rendered diffuse and, in practice, unenforceable without transnational agreements"[39]. National legal reforms are therefore insufficient without international harmonization. And without internal oversight, institutions may unknowingly place patients at risk by adopting tools whose real-world efficacy differs from promotional claims or lab results.

From a doctrinal perspective, the principle of primum non nocere (first, do no harm) has a link with constitutional protections of life and health. It is held that "legal systems must build AI governance not just around safety and efficacy, but around dignity, redress, and fairness"[40]. A purely instrumental view of AI as a technical tool to increase efficiency fails to consider the broader constitutional obligations to protect life, especially in nations where healthcare is a fundamental right. Positive law also needs to address structural inequalities worsened by biased AI systems. Courts should allow disparate impact claims against institutions that deploy biased AI tools, thereby extending equal protection guarantees into the algorithmic realm. Safeguarding human life in the AI-driven diagnostic era requires a shift in positive law from physician-centred liability to multi-actor responsibility that includes developers, institutions, and regulators. Legal frameworks must evolve from treating AI as a neutral tool to recognizing it as a system with agency-like effects, capable of transforming clinical judgment and impacting human rights.

Findings

A primary finding is that the application of Artificial Intelligence (AI) in medical diagnostics reveals a paradox between innovation and vulnerability as AI systems introduce a new category of error, distinct from traditional clinical mistakes. These are not necessarily due to lack of knowledge or negligence,

but are systemic stemming from limitations in data representativeness, model assumptions, and learning drift. This suggests that AI's touted objectivity is only as robust as its developmental context. A system trained on homogenous datasets cannot generalize ethically to diverse populations, thus undermining the equity AI is often assumed to promote. This becomes dangerous when clinicians over-rely on AI outputs, even in the face of contradictory patient presentations.

Again, there is a gap in positive law where causation is now shared between human and machine actors meaning legal liability cannot be assessed using binary standards of negligence. Here, AI systems may unintentionally erode informed consent and patient autonomy. The use of opaque AI systems, particularly those without explainable features, can obstructs a patient's right to meaningful choice. Moreover, even when disclosure occurs, comprehension may not follow. Informed consent must be comprehensible otherwise, it becomes procedural rather than participatory.

Discussion

Clinicians and institutions must be trained to bridge the gap above not merely in disclosing the use of AI but contextualize its role in lay-accessible language. Again, there is the aspect of the erosion of trust in clinical relationships when AI errors are revealed. This mistrust extends beyond the specific case as it generalizes to the institution and to AI tools more broadly. Importantly, trust is not only a psychological state but an ethical currency upon which consent, cooperation, and compliance depend. If AI jeopardizes this trust, it risks diminishing the effectiveness of medical care itself. Institutional gaps in AI governance, particularly regarding auditability and performance monitoring has to be bridged. In effect, patients should not be exposed to tools whose performance is not transparently monitored after deployment, a breach of both ethical due diligence and legal duty of care.

Additionally, there is the disconnect between regulatory aspiration and implementation. Most countries have drafted ethical principles on AI in healthcare transparency, fairness, and accountability but few have codified them into enforceable law. Ethical declarations without regulatory muscle therefore become symbolic rather than substantive. This is why diagnostic AI systems are not neutral tools. Their construction embeds normative assumptions, often invisible to end users. For instance, a system that minimizes false positives for cancer may increase false negatives, delaying critical treatment. Ethical and legal analyses must therefore interrogate the value priorities that shape algorithmic development.

AI systems in diagnosis amplify structural inequalities, often replicating or worsening existing disparities. Such disparities are not accidental as they are designed into the system via exclusion. Patients in low-resource settings or with rare genetic profiles are effectively rendered invisible to models built on majority data. This is not just a technical flaw but a distributive injustice with life-altering consequences. The implication is that AI, left unchecked, will stratify health outcomes along socio-demographic lines. Positive laws must explicitly mandate inclusivity in training data and performance reporting not aggregated by race, gender, and socio-economic background.

Conclusion

The ethical and legal implications of AI in diagnostic medicine are no longer speculative but current and fully documented, revealing that technological deployment without ethical architecture exposes patients to harm, clinicians to unfair liability, and institutions to governance failure. This study has shown that while AI holds transformative potential for improving diagnostic accuracy and efficiency, its integration into healthcare introduces significant ethical dilemmas and legal uncertainties. AI tools may promise objectivity but are deeply contingent on subjective design choices, biased training data, and opaque decision-making structures.

More broadly, AI challenges the ontological assumptions of medical care as diagnosis ceases to be a dialogic, interpretive act but becomes a statistical calculation. This shift not only compresses patient identity into data but erodes the moral presence of the clinician. To restore moral and legal clarity, technological innovation must be aligned with ethical discernment and supported by enforceable legal norms. Arising from these, regulatory authorities must require all AI systems used in diagnostic processes to include explainable features. Legal reforms should also move toward

shared or strict liability frameworks in which responsibility is distributed among developers, deployers, and clinicians.

Recommendations

I. Healthcare institutions must establish internal AI ethics committees as a matter of urgency to conduct ongoing audits of AI performance, fairness, and patient impact. These audits should be transparent, independent, and responsive to real-world deployment outcomes.
ii. Medical education must also integrate AI ethics and critical AI literacy to ensure clinicians understand both the capabilities and limitations of diagnostic systems.
iii. Laws and guidelines must require that AI systems be trained and tested on diverse datasets that reflect the demographic, genetic, and cultural diversity of the populations they serve just as AI governance should prioritize patient safety, dignity, and rights.
iv. Lastly, to ensure just outcomes when diagnostic harm occurs, governments or insurers should establish no-fault compensation pools funded by AI vendors and health systems. As it is only through rigorous governance, an interdisciplinary engagement such as this effort and a steadfast commitment to human dignity can AI in medical diagnostics serve and save life rather than merely calculate it.

References

1. Carter, S. M., Rogers, W., Win, K. T., Frazer, H., & Richards, B. (2020). The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *The Breast*, 49,20–25.
2. Kothinti, R. R. (2024). Artificial intelligence in healthcare. *World Journal of Advanced Research and, Reviews*,20(3),118–126.
3. Chikhaoui, E., Alajmi, A., & Shaikh, S. (2022). Artificial intelligence applications in healthcare: Ethical and legal challenges. *Emerging Science Journal*, 6(2).
4. Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., & Shah, M. (2022). Legal and ethical considerations in AI healthcare. *Frontiers in Surgery*, 9, Article 862322.
5. Pesapane, F., Volonté, C., Codari, M., & Sardanelli, F. (2018). Artificial intelligence as a medical device in radiology: Ethical and regulatory issues in Europe and the United States. *Insights into Imaging*, 9(5), 745–753.
6. Schönberger, D. (2019). AI in healthcare: Legal and ethical implications. *International Journal of Law and Information Technology*, 27(2), 171–203.
7. Kothinti, R. R. (2024). Artificial intelligence in healthcare. *World Journal of Advanced Research and, Reviews*,20(3),118–126.
8. Gooding, P., & Kariotis, T. (2021). Ethics and law in research on algorithmic and data-driven technology in mental health care: Scoping review. *JMIR Mental Health*, 8(6), e24668.
9. World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. WHO.
10. Amann, J., Blasimme, A., Vayena, E., & Frey, D. (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20(1), 310.
11. Warrell, D.A, Cox T.M., Firth, J.D, & Benz, E.J. (Eds) (2010). *Oxford textbook of medicine* (5th ed., Vol 1).Oxfrod University Press.
12. Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and legal challenges of artificial intelligence–driven healthcare. In D. D. Bikson (Ed.), *Artificial Intelligence in Healthcare Elsevier* 295–336.
13. Archibong, Emmanuel Iniobong (2024) Beautiful Metaphysics: *A Western and African Tapestry*. Robert Minder Publisher.
14. Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., & Shah, M. (2022). Legal and ethical considerations in AI healthcare. *Frontiers in Surgery*, 9, Article 862322.
15. World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. WHO.
16. Schönberger, D. (2019). AI in healthcare: Legal and ethical implications. *International Journal of Law and Information Technology*, 27(2), 171–203.
17. Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., & Shah, M. (2022). Legal and ethical considerations in AI healthcare. *Frontiers in*

Surgery, 9, Article 862322.
18. Schönberger, D. (2019). AI in healthcare: Legal and ethical implications. *International Journal of Law and Information Technology*, 27(2), 171–203.
19. Refolo, P., Raimondi, C., Astratinei, V., & Battaglia, L. (2025). Assessment of AI-based technologies in HTA. *Journal of Medical Internet Research*, 27(1), e11988925.
20. Peičius, E., Astromskis, P., & Astromskė, K. (2021). Ethical and legal challenges of informed consent. *AI & Society*, 36(1), 203–210.
21. Kothinti, R. R. (2024). Artificial intelligence in healthcare. *World Journal of Advanced Research and, Reviews*,20(3),118–126.
22. Tang, L., Li, J., & Fantus, S. (2023). Medical artificial intelligence ethics: A systematic review of empirical studies. *Digital Health*, 9, 20552076231186064.
23. Abdullah, Y. I., & Schuman, J. S. (2021). Ethics of artificial intelligence in medicine and ophthalmology. *Asia-Pacific Journal of Ophthalmology*, 10(2), 43–48.
24. Mahomed, S. (2018). Healthcare, AI and the Fourth Industrial Revolution. *South African Journal of Bioethics and Law*, 11(2).
25. Carter, S. M., Rogers, W., Win, K. T., Frazer, H., & Richards, B. (2020). The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *The Breast*, 49,20–25.
26. Čartolovni, A., Tomičić, A., & Mosler, E. L. (2022). Ethical, legal, and social considerations of AI-based medical decision-support tools. *International Journal of Medical Informatics*, 163.
27. Archibong, Emmanuel Iniobong (2024) "The Ontology of the minds creativity and the language of artificial Intelligence". USEM Journal of Languages, Linguistics and Literature. Vol 12,
28. Zhang, J., & Zhang, Z. (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC Medical Informatics and Decision Making*, 23(1), 1–9.
29. Ibrahim, S. A., & Pronovost, P. J. (2021). Diagnostic errors, health disparities, and artificial intelligence: A combination for health or harm? JAMA Health Forum, 2(9), e213129. https://doi.org/10.1001/jamahealthforum.2021.3129
30. Mensah, G. B. (2024). Ethics, bias, and risks in AI diagnosis. *ResearchGate*. https://www.researchgate.net/publication/381584371
31. Gozum, I. E. A., & Flake, C. C. D. (2024). Human dignity and artificial intelligence in healthcare. *Journal of Religion and Health*.
32. Love, C. S. (2023). "Just the facts, ma'am": Ethical considerations for AI in medicine and its impact on hope. The Linacre Quarterly, 90(1), 1 – 8 . https://doi.org/10.1177/00243639231162431
33. Kunnathu, J. (2024). An ethical evaluation of artificial intelligence from a Catholic moral perspective. Asian Horizons, 18(2), 9–19. https://dvkjournals.in/index.php/ah/article/view/4592
34. Török, B., & Darabos, Á. (2025). EU tech regulation and Catholic teaching on AI. *Religions,* 16(3), 312.
35. Ghosh, D. P. (2024). Moral gravity in engineered empathy. *ResearchGate*.
36. Smith, H., & Fotheringham, K. (2020). AI in clinical decision-making. *Medical Law International*, 20(3), 211–226.
37. Bottomley, D., & Thaldar, D. (2023). Liability for harm caused by AI in healthcare: An overview of the core legal concepts. *Frontiers in Pharmacology*, 14.
38. World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. WHO.
39. Chau, M. (2024). Ethical, legal, and regulatory landscape of artificial intelligence in Australian healthcare. *Journal of Medical Imaging and Radiation Sciences*, 55(1), 110–115.
40. Zhang, J., & Zhang, Z. (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC Medical Informatics and Decision Making*, 23(1), 1–9.