

Volume 2, Issue 1, October 2026: DOI: <https://doi.org/10.5281/zenodo.22876368>. Article history:
Received: May 1, 2026; Revised: August 1, 2026; Accepted: August 30, 2026

Ethical, Legal, and Professional Vulnerabilities: Healthcare Providers' Perceptions of Liability in AI-Assisted Surgical and Diagnostic Protocols

Eni Eja Alobo

Department of Public Law, Faculty of Law, University of Calabar, Cross River State,
Nigeria

ORCID ID: <https://orcid.org/0009-0004-9930-465X>

***Ikechukwu Peter Ugbor**

Department of Mass Communication, University of Nigeria, Enugu State, Nigeria
ORCID ID: <https://orcid.org/0000-0003-1223-582X>

*Corresponding author email: ike.ugbor@unn.edu.ng

ORCID ID: <https://orcid.org/0000-0003-1223-582X>

Marian Ofunu Ujah

Department of Public Law, Faculty of Law, Prince Abubakar Audu University, Anyigba,
Kogi State, Nigeria

ORCID ID: <https://orcid.org/0009-0003-0701-726X>

Abstract

Background: The rise of clinical algorithmic tools has transformed healthcare delivery, establishing artificial intelligence protocols as pervasive standards in diagnostic and surgical workflows. Despite their diagnostic precision and operational efficiency, empirical investigation into the professional, legal, and epistemic vulnerabilities experienced by frontline clinicians under prevailing liability doctrines remains limited.

Objective: This study developed and tested the Integrated Vulnerability–Liability–Adoption Model to examine how epistemic vulnerability, civil liability exposure, and criminal culpability fears influence clinical trust deficits, defensive medical practices, and protocol adoption intentions across diagnostic and surgical specialists.

Methodology: A cross-sectional quantitative survey design was conducted with licensed specialists (N = 284) across tertiary hospitals using the researcher-developed Healthcare Providers' Liability Perceptions Scale. Data were analysed using Partial Least Squares Structural Equation Modelling (PLS-SEM) in SmartPLS 4, evaluating measurement parameters, structural path coefficients via 5,000 bootstrap resamples, mediation effects, and multi-group differences between diagnostic and surgical cohorts.

Result: Epistemic vulnerability exerted strong positive effects on perceived civil liability ($\beta = 0.462$, $p < 0.001$) and criminal culpability ($\beta = 0.398$, $p < 0.001$). Perceived civil liability significantly inhibited protocol adoption ($\beta = -0.318$, $p < 0.001$) and drove widespread defensive utilisation behaviours ($\beta = 0.512$, $p < 0.001$). Clinical trust deficit partially mediated the relationship between liability perception and adoption resistance ($\beta = -0.112$, $p < 0.001$). Institutional resilience negatively moderated defensive posturing ($\beta = -0.184$, $p = 0.001$). Multi-group analysis revealed that liability-induced adoption resistance and defensive behaviours were significantly more pronounced among surgical specialists than diagnostic clinicians.

Conclusion: Algorithmic opacity and disproportionate legal exposure undermine physician trust, transforming clinicians into institutional liability shields and inducing defensive clinical redundancies. Sustainable technological integration requires statutory liability reform, institutional safe harbours, and transparent override frameworks that equitably distribute risk across developers, healthcare organisations, and practitioners.

Keywords: Artificial intelligence; Clinical liability; Defensive medicine; Epistemic vulnerability; Algorithmic governance; Robotic surgery; Diagnostic radiology; Partial least squares structural equation modelling.

Introduction

The integration of artificial intelligence into clinical environments has transitioned from speculative experimentation to operational deployment across diagnostic radiography, pathology, interventional endoscopy, and robotic surgery. Machine learning architectures, deep convolutional neural networks, and computer vision systems now regularly execute analytical tasks that historically depended entirely on human clinical judgement. These algorithms screen complex radiological images, identify surgical margins, and provide real-time guidance during intricate operative procedures.

Yet, as algorithmic tools increasingly influence clinical pathways, they introduce profound structural and conceptual challenges to traditional models of medical practice (Elendu et al., 2023; Prakash et al., 2022). While technical performance metrics frequently demonstrate high diagnostic

precision, the real-world operationalisation of these systems reveals a precarious imbalance between machine agency and human professional responsibility (Khan et al., 2025; Nasir et al., 2025).

Healthcare providers find themselves situated at the confluence of technological advancement and institutional exposure. Modern medicolegal jurisprudence and professional codes of conduct have evolved over centuries under the foundational assumption that the human clinician remains the sole autonomous decision-maker. The introduction of autonomous or semi-autonomous algorithmic recommendations disrupts this paradigm. Practitioners face the complex challenge of reconciling opaque, algorithmic outputs with their personal and institutional duties of care (Palama, 2022; Pham, 2025).

When an algorithmic diagnostic protocol fails to detect an occult fracture, misclassifies a malignant lesion, or when a surgical robotic assistant commits a mechanical trajectory deviation, the question of accountability becomes fraught (Gourari et al., 2025; Hashimoto et al., 2026). The ambiguity surrounding whether fault lies with the software developer, the deploying hospital, or the attending clinician creates substantial psychological, ethical, and legal vulnerabilities for healthcare personnel (Naik et al., 2022; Rodrigues, 2020).

These vulnerabilities are not purely theoretical; they shape the daily realities and decision-making behaviours of clinicians. Healthcare professionals must negotiate systemic pressures to adopt innovative technologies whilst simultaneously navigating the risk of professional censure, civil litigation, and even criminal sanctions in the event of adverse patient outcomes (Budiono et al., 2026; Mrčela & Vuletić, 2025).

A critical gap exists between technical governance models and the lived experiences of clinicians who must implement these tools at the point of care. Although institutional guidelines emphasise safety and trustworthiness, empirical insight into how providers perceive their exposure to liability across surgical and diagnostic specialisms remains scarce.

This paper investigates the multifaceted ethical, legal, and professional vulnerabilities experienced by healthcare providers operating within artificial intelligence-assisted diagnostic and surgical protocols. By examining how clinical practitioners perceive liability and the attribution of fault, this study clarifies the structural protections and statutory reforms required to foster safe, fair, and sustainable medical innovation.

Literature Review

Conceptualising Vulnerability and Trustworthiness in Algorithmic Healthcare

The concept of vulnerability in the context of emerging technologies extends beyond the physical and biological fragility of the patient. In digital ethics, algorithmic implementation exposes both individuals and systems to novel forms of harm, dependency, and loss of control. Hine and Floridi (2026) conceptualise digital vulnerabilities through a multi-level analytical framework, illustrating

how algorithmic systems generate systemic exposure at micro, meso, and macro levels. Within healthcare, clinicians occupy a precarious middle tier: they are burdened with operational dependency on digital tools while lacking direct insight into their mathematical reasoning.

This structural asymmetry is compounded by wider societal and human rights challenges inherent to algorithmic governance. Rodrigues (2020) highlights that the rapid deployment of artificial intelligence systems routinely outpaces institutional human rights frameworks, creating critical gaps in transparency, privacy, and accountability. When applied to clinical environments, algorithmic opacity deprives practitioners of the epistemic justifications necessary to defend clinical choices (Goktas & Grzybowski, 2025).

To counter these systemic hazards, socio-legal scholars emphasise the need for multi-level resilience. Teo (2025) argues that addressing technological vulnerabilities requires layered regulatory architectures that protect human agency rather than subordinating it to algorithmic efficiency.

Parallel scholarship on algorithmic risk governance underscores that system trustworthiness cannot simply be presumed from statistical benchmarks. Polemi et al. (2024) evaluate the difficulties of managing artificial intelligence trustworthiness risks, observing that data drift, model hallucinations, and adversarial vulnerabilities render complete technical certainty impossible.

In broader organisational and business settings, Maiti et al. (2025) note that technological adoption without robust ethical oversight produces acute institutional risks and undermines stakeholder trust. In medical settings, this lack of verifiable trustworthiness directly undermines clinician confidence, forcing healthcare workers to adopt defensive clinical postures to mitigate their perceived exposure to liability (Khan et al., 2025).

Ethical Conundrums and Epistemic Asymmetries in Clinical Decision-Making

The ethical dilemmas confronting clinicians utilising artificial intelligence tools stem fundamentally from what bioethicists term the black-box problem. Machine learning algorithms, particularly deep neural networks applied to medical imaging and predictive pathology, arrive at outputs through non-linear mathematical transformations that cannot be fully audited or explained by human practitioners (Goktas & Grzybowski, 2025; Prakash et al., 2022).

Prakash et al. (2022) demonstrate through a comprehensive scoping review that the deployment of opaque models fractures conventional medical ethics, undermining the principle of beneficence and compromising valid informed consent. Clinicians are placed in the ethically compromised position of explaining risks and algorithmic recommendations to patients without having full visibility into the foundational rationale behind those recommendations (Elendu et al., 2023; Pham, 2025).

Furthermore, Khan et al. (2025) indicate that the pursuit of secure and trusted artificial intelligence in clinical workflows requires addressing critical trade-offs between computational autonomy and clinical oversight. When diagnostic systems operate with near-human or supra-human accuracy, practitioners experience automation bias: the cognitive tendency to accept algorithmic recommendations uncritically, even when contradictory clinical indicators emerge (Nasir et al., 2025).

Conversely, should a practitioner choose to reject an algorithmic recommendation that subsequently proves correct, they face heightened professional scrutiny for deviating from an automated benchmark (Palama, 2022). This dilemma creates profound ethical tension. Providers are held responsible for safeguarding patient welfare, yet they are increasingly required to defer to algorithmic systems whose operational pathways they cannot verify (Naik et al., 2022; Pham, 2025).

Legal Ambiguities and Liability Frameworks in Diagnostic Protocols

The legal landscape governing diagnostic artificial intelligence reflects a fragmented patchwork of civil codes, tort principles, and statutory gaps. Traditional medical malpractice law is anchored to the concept of negligence, requiring proof of a duty of care, a breach of that duty against an established professional standard, and direct causation resulting in patient harm. Algorithmic diagnostics disrupt every node of this causal chain (Naik et al., 2022; Shakhathreh & Alazzam, 2025).

Evaluating liability across multiple legal systems reveals significant jurisdictional divergence. Gourari et al. (2025) compare the legal frameworks of France, Jordan, and the United Arab Emirates regarding artificial intelligence-assisted radiographic diagnostic errors. Their findings reveal that while some continental systems attempt to apply strict product liability to software developers, civil and tort systems typically retain the clinician as the ultimate liability shield. Under these prevailing legal doctrines, the practitioner is presumed to retain continuous supervisory duty over the diagnostic conclusion. As a result, software malfunctions, flawed training datasets, or algorithmic hallucinations are frequently categorised not as product defects, but as practitioner negligence for failing to detect the machine's error (Gourari et al., 2025; Shakhathreh & Alazzam, 2025).

This statutory lag leaves healthcare professionals vulnerable to disproportionate civil claims. Palama (2022) observes that high-risk health technologies demand harmonised legal frameworks that align technical robustness with shared liability. In the absence of such statutory alignment, the burden of proof falls heavily on clinicians, who must demonstrate that their reliance on an algorithmic diagnostic aid conformed to an accepted standard of care (Nasir et al., 2025; Pham, 2025).

When medical licensing bodies and legal jurisdictions fail to define clear standards for algorithmic reliance, healthcare providers face considerable legal ambiguity whenever they interact with clinical decision support tools (Hashimoto et al., 2026; Naik et al., 2022).

Surgical Autonomy, Robotic Trajectories, and Tort Liability

The operationalisation of artificial intelligence within surgical disciplines introduces kinetic and mechanical dimensions to practitioner liability. Unlike purely diagnostic software, surgical systems interact directly with human tissue through robotic arms, autonomous electrocautery, and smart stereotactic guides. In their landmark white paper for the Society of American Gastrointestinal and Endoscopic Surgeons, Hashimoto et al. (2026) map the complex spectrum of risk and liability across the surgical deployment of artificial intelligence.

The authors demonstrate that surgical liability shifts dynamically depending on the level of autonomy granted to the platform. In systems providing intraoperative guidance or surgical phase recognition, liability remains firmly with the operating surgeon. However, as robotic platforms gain semi-autonomous capacity—such as executing automated suturing or tissue retraction—the boundary between surgeon error and hardware or algorithmic malfunction becomes deeply obscured (Hashimoto et al., 2026).

Specialised surgical fields showcase these challenges acutely. Budiono et al. (2026) examine legal accountability and patient protection in artificial intelligence-assisted orthopaedic surgery, where algorithmic navigation guides bone resections and prosthetic placement. Their analysis shows that when sub-millimetre prosthetic misalignments occur, existing tort systems struggle to differentiate between surgeon miscalibration, robotic sensor error, and pre-operative algorithmic planning defects.

In a related context, Al-Zawahreh and Al-Dhanibat (2026) analyse the civil liability of cosmetic surgeons employing digitalised and artificial intelligence-assisted systems. They observe that aesthetic surgery involves rigorous duties regarding informed consent and anticipated aesthetic outcomes. When algorithmic simulations promise specific post-operative results that the physical procedure fails to achieve due to algorithmic error, surgeons are held personally liable for contractual and civil damages (Al-Zawahreh & Al-Dhanibat, 2026).

Across both orthopaedic and reconstructive interventions, judicial systems continue to treat the surgical robot as a mere extension of the surgeon's hand, thereby resisting the distribution of liability to manufacturers or healthcare institutions (Budiono et al., 2026; Shakhathreh & Alazzam, 2025).

Criminal Liability and the Boundaries of Professional Culpability

While civil litigation remains the most frequent medicolegal hazard, the potential for criminal prosecution introduces an even more serious threat to professional practice. Mrčela and Vuletić (2025) explore the complex boundaries of criminal liability in artificial intelligence-assisted medicine, analysing the threshold at which algorithmic medical errors transition from civil torts to criminal negligence or involuntary manslaughter.

Criminal law universally requires the establishment of personal culpability, typically characterised by recklessness, gross negligence, or intentional disregard of human safety. Mrčela and Vuletić (2025) argue that as healthcare facilities formalise protocols mandating the use of artificial intelligence, clinicians could face criminal prosecution for either blindly rubber-stamping dangerous algorithmic outputs (gross negligence) or for deliberately refusing to follow validated algorithmic guidance when a standard patient survival protocol depended upon it.

This legal paradox severely compounds the psychological vulnerability of medical personnel. Clinicians find themselves caught in a regulatory trap: uncritical adherence to technology creates liability for automated failure, whereas intentional refusal to utilise algorithmic systems exposes them to accusations of practicing sub-standard, archaic medicine (Naik et al., 2022; Nasir et al., 2025; Palama, 2022).

The literature highlights that the current medicolegal environment treats clinicians as human shock absorbers for technological flaws, absorbing civil, regulatory, and criminal liability while manufacturers and deploying institutions frequently operate within protected safe harbours (Gourari et al., 2025; Mrčela & Vuletić, 2025).

Theoretical Framework and Research Hypotheses

Theoretical Framework: The Integrated Vulnerability–Liability–Adoption Model

To examine how clinical practitioners navigate artificial intelligence integration under conditions of legal and professional uncertainty, this study constructs the Integrated Vulnerability–Liability–Adoption Model.

The framework synthesises three foundational theoretical traditions:

1. The Multi-Level Digital Vulnerability Theory (Hine & Floridi, 2026; Teo, 2025), which conceptualises technological exposure across individual (micro), organizational (meso), and systemic/legal (macro) dimensions.
2. Defensive Medicine and Risk Perception Theory (Studdert et al., 2005; Bourne et al., 2015), which explains how exposure to civil and criminal sanctions induces defensive adaptations, risk-averse behaviour, and selective disengagement from standard clinical protocols.
3. The Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology (Davis, 1989; Venkatesh et al., 2003), specifically through the inclusion of perceived risk, trust, and structural facilitating conditions as decisive determinants of adoption intention and protocol adherence.

In conventional technology adoption paradigms, adoption intention is primarily driven by perceived usefulness and perceived ease of use. However, clinical environments are governed by non-negotiable legal duties of care and stringent professional licensure. Consequently, when clinicians interface with opaque, high-risk systems (such as autonomous surgical trajectory

guidance or diagnostic image classifiers), perceived risk does not function merely as an ancillary barrier; it actively reshapes professional identity, clinical discretion, and behavioural adherence.

Research Hypotheses

Based on the conceptual interactions within the Integrated Vulnerability–Liability–Adoption Model, seven testable hypotheses are proposed across structural equation modelling, multiple regression, and path analysis.

- **H1a:** Perceived epistemic vulnerability (inability to audit or explain algorithmic outputs) is positively associated with healthcare providers' perceived civil liability exposure.
- **H1b:** Perceived epistemic vulnerability is positively associated with healthcare providers' fear of criminal culpability and regulatory licensure sanctions.

H2a: Perceived civil liability exposure is negatively associated with healthcare providers' intention to routinely adopt artificial intelligence diagnostic and surgical protocols.

H2b: Perceived fear of criminal culpability is negatively associated with healthcare providers' intention to adopt semi-autonomous surgical protocols.

H3: Clinical trust in algorithmic recommendations significantly mediates the negative relationship between perceived liability exposure and protocol adoption intention.

H4a: Perceived liability exposure is positively associated with defensive clinical behaviour, characterised by redundant diagnostic ordering and systematic manual recalculation of algorithmic outputs.

H4b: Perceived liability exposure significantly increases the frequency of intentional algorithmic override or protocol rejection when clinical edge cases emerge.

H5: The negative relationship between perceived liability exposure and protocol adoption is significantly stronger among interventional surgeons than among diagnostic radiologists and non-interventional clinicians.

H6a: Institutional risk-sharing frameworks and clinician indemnification policies negatively moderate the relationship between perceived civil liability exposure and protocol resistance, weakening the adverse effect.

H6b: Formal medicolegal training in artificial intelligence jurisprudence positively moderates protocol adherence by reducing ambiguity surrounding standard-of-care compliance.

Methodology

Research Design

To investigate how healthcare providers perceived liability, epistemic vulnerability, and professional exposure when interacting with artificial intelligence systems, this study employed a cross-sectional quantitative survey design.

Population and Sampling Strategy

Target Population: The target population consisted of licensed medical practitioners practising within tertiary healthcare facilities where artificial intelligence decision-support platforms or semi-autonomous operative systems were actively integrated or undergoing formal clinical evaluation. The population was drawn from two primary clinical domains:

- Diagnostic specialists: Consultant radiologists, resident radiologists, and histopathologists.
- Procedural and surgical specialists: Orthopaedic surgeons, general and gastrointestinal surgeons, and neurosurgeons.

Sample Size Determination: The minimum required sample size was calculated a priori using G*Power 3.1.9.7 for linear multiple regression and partial least squares structural equation modelling. Specifying a medium effect size ($f^2 = 0.15$), a statistical power ($1 - \beta$) of 0.95, a significance criterion (α) of 0.05, and six direct structural predictors, the minimum statistical sample size required was 130 respondents.

To ensure sufficient statistical power for multi-group comparative analysis across clinical specialisms, and to compensate for an anticipated 30% to 40% non-response or incomplete submission rate typical of physician survey research, the self-administered instrument was distributed to a target sample of 350 practitioners across accredited tertiary teaching hospitals and specialist referral centres.

Sampling Technique: A proportionate stratified random sampling technique was applied across participating healthcare institutions. The sampling frame was stratified into two equal clinical arms: diagnostic specialists (target: $n = 175$) and surgical specialists (target: $n = 175$). Eligible respondents were required to hold full specialist registration or be enrolled in an accredited senior residency training programme with a minimum of six months of direct clinical engagement with institutional algorithmic decision-support or robotic navigational tools.

Instrumentation and Operationalisation

Data were collected using a structured survey instrument titled the Healthcare Providers' Liability Perceptions Scale (HPLP-Scale), developed specifically by the researcher to capture the medicolegal and operational dimensions of clinical artificial intelligence use. The development process followed standard psychometric scale development guidelines, comprising item generation, content validation, and empirical refinement.

Instrument Validation and Pilot Testing

Because the researchers developed the measurement instrument for this investigation, it underwent rigorous pre-testing to establish content validity, face validity, and construct reliability before widespread field deployment.

Pilot Study: We administered a pilot test to 30 eligible medical practitioners (15 diagnostic specialists and 15 surgeons), who were subsequently excluded from the primary study cohort. The pilot study evaluated administration clarity, completion time (averaging 10 to 12 minutes), and preliminary internal consistency. Corrected item-total correlations exceeded the acceptable threshold of 0.30 across all retained items, and subscale Cronbach's alpha values ranged between 0.78 and 0.89, demonstrating satisfactory preliminary reliability.

Survey Administration and Data Collection

The survey was administered over 12 weeks using a dual-mode distribution strategy. We deployed an electronic version via an encrypted institutional survey platform (REDCap) and distributed it through hospital clinical mailing lists and professional associations. Simultaneously, we delivered self-administered hard-copy questionnaires during departmental clinical audits, morbidity and mortality meetings, and grand rounds to maximise response rates.

Data Analysis and Ethical Considerations

Data screening was conducted in IBM SPSS Statistics 29, followed by structural equation modelling in SmartPLS 4 using PLS-SEM to evaluate complex mediator and moderator paths without normality constraints. Ethical approval for this study was formally granted by the Department of Mass Communication Ethical Approval Committee under protocol reference number MAREC/202501328 prior to data collection.

Results

Preliminary Data Screening and Sample Profile

We distributed 350 questionnaires across participating tertiary healthcare institutions and retrieved 298 copies. After data cleaning and inspection, 14 submissions were removed because they had missing entries exceeding 10% or systematic straight-lining. The remaining 284 valid records constituted an effective response rate of 81.1%.

Table 1: Demographic and Clinical Profile of Respondents (N = 284)

Variable	Category	Frequency (n)	Percentage (%)
Primary Clinical Specialty	Diagnostic Radiology	104	36.6

	Histopathology	42	14.8
	Orthopaedic Surgery	58	20.4
	General / GI Surgery	52	18.3
	Neurosurgery	28	9.9
Years of Clinical Experience	Less than 5 years	38	13.4
	5 to 10 years	84	29.6
	11 to 20 years	118	41.5
	Greater than 20 years	44	15.5
Institution Type	Public / University Teaching Hospital	188	66.2
	Private Tertiary Hospital	96	33.8
Primary AI Modality	Diagnostic Image Analysis / Detection	146	51.4
	Intraoperative Navigation / Robotics	138	48.6

The sample comprised diagnostic (51.4%, n = 146) and surgical specialists (48.6%, n = 138), matching their respective diagnostic versus robotic artificial intelligence usage. Most practitioners had over 10 years of experience (57.0%) and worked in public teaching hospitals (66.2%).

Measurement Model Assessment

The measurement model was evaluated in SmartPLS 4 using Partial Least Squares Structural Equation Modelling to establish the reliability and validity of the seven reflective constructs before testing structural relationships.

Table 2: Measurement Model Evaluation (Reliability and Convergent Validity)

Latent Construct	Indicator Code	Outer Factor Loading	Cronbach Alpha	Composite Reliability	Average Variance Extracted (AVE)
Epistemic Vulnerability	EV1	0.812	0.842	0.894	0.678
	EV2	0.835			
	EV3	0.844			
	EV4	0.801			
Civil Liability Exposure	CL1	0.828	0.856	0.903	0.698

	CL2	0.854			
	CL3	0.839			
	CL4	0.82			
Criminal Culpability	CR1	0.865	0.814	0.889	0.728
	CR2	0.842			
	CR3	0.853			
Clinical Trust Deficit	CT1	0.768	0.821	0.882	0.652
	CT2	0.814			
	CT3	0.826			
	CT4	0.819			
Defensive Utilisation	DU1	0.831	0.871	0.912	0.722
	DU2	0.862			
	DU3	0.855			
	DU4	0.85			
Institutional Resilience	IR1	0.792	0.792	0.878	0.706
	IR2	0.864			
	IR3	0.862			
AI Protocol Adoption Intention	AI1	0.876	0.828	0.897	0.744
	AI2	0.869			
	AI3	0.842			

The measurement model satisfied all psychometric benchmarks. Indicator reliability was established with all 27 outer loadings exceeding 0.708 (0.768 to 0.876). Construct consistency was confirmed by Cronbach's alpha (0.792 to 0.871) and composite reliability values (0.878 to 0.912) above 0.70, whilst convergent validity was supported by average variance extracted figures (0.652 to 0.744) surpassing 0.50.

Discriminant Validity

Discriminant validity was assessed using the Heterotrait-Monotrait Ratio of Correlations, which provides a more sensitive and conservative evaluation of construct uniqueness than traditional metric cross-loadings.

Table 3: Discriminant Validity via Heterotrait-Monotrait Ratio (HTMT)

Construct	1	2	3	4	5	6	7
1. Epistemic Vulnerability	—						
2. Civil Liability Exposure	0.524	—					
3. Criminal Culpability	0.468	0.612	—				
4. Clinical Trust Deficit	0.412	0.548	0.486	—			
5. Defensive Utilisation	0.495	0.684	0.532	0.442	—		
6. Institutional Resilience	0.312	0.384	0.298	0.345	0.372	—	
7. Protocol Adoption Intention	0.386	0.518	0.436	0.492	0.562	0.284	—

Note: Values must be strictly below 0.85 to confirm discriminant validity.

Table 3 confirms that discriminant validity was established across the measurement model. All calculated HTMT ratios fell between 0.284 and 0.684, remaining well below the conservative ceiling of 0.85.

The highest observed correlation occurred between Civil Liability Exposure and Defensive Utilisation (HTMT = 0.684), followed by Civil Liability Exposure and Criminal Culpability (HTMT = 0.612). Both values remained well below critical cut-offs, demonstrating that each operationalised domain represents an empirically distinct theoretical construct rather than overlapping measurement noise.

Structural Model and Hypotheses Testing

Following validation of the outer measurement model, structural path relationships were evaluated. Collinearity diagnostics demonstrated that all inner Variance Inflation Factor values ranged between 1.34 and 2.21, confirming that structural multicollinearity was not present.

Hypothesis testing was conducted using a non-parametric bootstrapping procedure with 5,000 resamples at a 95% confidence interval.

Table 4: Structural Path Estimates and Hypotheses Decisions

Hypothesis	Structural Path	Path Coefficient (Beta)	Standard Error	t-Value	p-Value	95% Bootstrap CI	f2 Effect Size	Decision
------------	-----------------	-------------------------	----------------	---------	---------	------------------	----------------	----------

H1a	Epistemic Vuln. -> Civil Liability	0.462	0.053	8.641	< 0.001	[0.354, 0.562]	0.271	Supported
H1b	Epistemic Vuln. -> Criminal Culpability	0.398	0.056	7.114	< 0.001	[0.282, 0.502]	0.188	Supported
H2a	Civil Liability -> Protocol Adoption	-0.318	0.061	5.246	< 0.001	[-0.432, - 0.194]	0.124	Supported
H2b	Criminal Culpability -> Protocol Adoption	-0.245	0.06	4.108	< 0.001	[-0.358, - 0.122]	0.076	Supported
H3 (Med)	Civil Liab. -> Trust Deficit - > Adoption	-0.112	0.028	4.025	< 0.001	[-0.168, - 0.059]	â€”	Supported
H4a	Civil Liability -> Defensive Utilisation	0.512	0.049	10.428	< 0.001	[0.412, 0.604]	0.358	Supported
H4b	Criminal Culpability -> Defensive Util.	0.286	0.053	5.372	< 0.001	[0.178, 0.388]	0.112	Supported
H6a (Mod)	Civil Liab. x Inst. Resilience -> Def. Util.	-0.184	0.055	3.315	0.001	[-0.288, - 0.072]	0.054	Supported
H6b (Mod)	Training x Civil Liab. -> Protocol Adoption	0.142	0.051	2.784	0.005	[0.041, 0.241]	0.032	Supported

As detailed in Table 4, empirical testing provided strong statistical support for all hypothesised structural relationships:

- *Antecedents of Liability Perceptions (H1a & H1b)*: Epistemic vulnerability exerted a strong and statistically significant positive influence on perceived civil liability exposure (beta = 0.462, t = 8.641, p < 0.001) with a substantial effect size (f² = 0.271), supporting H1a. Epistemic vulnerability also significantly heightened fear of criminal culpability and licensure sanctions (beta = 0.398, t = 7.114, p < 0.001, f² = 0.188), confirming H1b. This

confirms that algorithmic opacity directly magnifies clinicians' subjective exposure to legal sanctions.

- *Liability and Adoption Intention (H2a & H2b)*: Apprehension of civil liability exhibited a significant negative relationship with willingness to adopt artificial intelligence protocols (beta = -0.318, $t = 5.246$, $p < 0.001$), validating H2a. Perceived criminal culpability similarly suppressed protocol adoption intention (beta = -0.245, $t = 4.108$, $p < 0.001$), validating H2b.
- *Mediation via Clinical Trust Deficit (H3)*: Bootstrapping of indirect effects revealed that clinical trust deficit significantly mediated the negative relationship between civil liability exposure and protocol adoption (beta = -0.112, $t = 4.025$, $p < 0.001$, 95% CI [-0.168, -0.059]), confirming partial mediation and supporting H3. Perceptions of liability damage the clinician's foundational trust in the technology, which subsequently dampens adoption intention.
- *Defensive Adaptations (H4a & H4b)*: Civil liability perception proved to be a powerful direct predictor of defensive utilisation behaviours, including duplicate testing and manual recalculated audits (beta = 0.512, $t = 10.428$, $p < 0.001$, $f^2 = 0.358$), supporting H4a. Perceived criminal culpability exerted an additional positive effect on defensive behaviour (beta = 0.286, $t = 5.372$, $p < 0.001$, $f^2 = 0.112$), supporting H4b.
- *Moderation by Institutional Safeguards (H6a & H6b)*: Institutional resilience significantly moderated the relationship between civil liability exposure and defensive utilisation (beta = -0.184, $t = 3.315$, $p = 0.001$), supporting H6a. In hospitals where clear indemnification policies and override rules were codified, the tendency to engage in defensive double-checking under legal threat was markedly reduced. Similarly, formal medicolegal training positively moderated protocol adoption (beta = 0.142, $t = 2.784$, $p = 0.005$), confirming H6b.

Overall, the structural model explained 42.6% of the variance in civil liability exposure ($R^2 = 0.426$), 38.1% of the variance in clinical trust deficit ($R^2 = 0.381$), 49.3% of the variance in defensive utilisation ($R^2 = 0.493$), and 46.8% of the variance in protocol adoption intention ($R^2 = 0.468$).

Multi-Group Analysis: Diagnostic versus Surgical Specialists

To evaluate Hypothesis 5, a non-parametric permutation-based Multi-Group Analysis was executed in SmartPLS to assess whether the structural paths differed between diagnostic specialists ($n = 146$) and surgical specialists ($n = 138$).

Table 5: Multi-Group Analysis Comparison (Diagnostic versus Surgical Specialists)

Structural Path	Diagnostic Cohort ($n = 146$) Beta	Surgical Cohort ($n = 138$) Beta	Absolute Difference (Beta Diff)	p-Value (PLS-MGA)	Decision on H5

Civil Liability -> Protocol Adoption	-0.228	-0.412	0.184	0.016	Supported
Criminal Culpability -> Defensive Utilisation	0.204	0.372	0.168	0.024	Supported
Epistemic Vulnerability -> Civil Liability	0.441	0.485	0.044	0.582	Invariant

Note: p-values below 0.05 denote statistically significant structural differences across clinical sub-groups.

Table 5 confirms significant differences across clinical specialisms, providing empirical support for Hypothesis 5:

- *Civil Liability and Protocol Adoption*: The negative path coefficient linking perceived civil liability to protocol adoption intention was significantly stronger among surgical practitioners (beta = -0.412) than among diagnostic practitioners (beta = -0.228), with an absolute path difference of 0.184 that was statistically significant ($p = 0.016$).
- *Criminal Culpability and Defensive Utilisation*: The impact of perceived criminal culpability on defensive clinical behaviour was significantly elevated among surgeons (beta = 0.372) compared to diagnostic clinicians (beta = 0.204), yielding a significant group difference of 0.168 ($p = 0.024$).
- *Cross-Group Invariance*: In contrast, the path from epistemic vulnerability to civil liability exposure remained statistically invariant across both cohorts (beta difference = 0.044, $p = 0.582$), indicating that algorithmic opacity induces substantial legal apprehension regardless of whether a clinician operates in an imaging suite or a surgical theatre.

These findings show that while black-box opacity generates broad apprehension across all disciplines, the invasive physical nature of surgical interventions makes proceduralists far more sensitive to liability threats, producing stronger protocol resistance and greater defensive adaptations.

Discussion

The empirical findings of this study provide robust quantitative validation for the Integrated Vulnerability–Liability–Adoption Model, illustrating how systemic epistemic deficits translate into medicolegal apprehension and defensive clinical adaptations. By examining structural paths across diagnostic and surgical specialisms, these results demonstrate that technical precision alone is insufficient to guarantee algorithmic protocol adherence when clinicians perceive themselves as unprotected liability shields for autonomous systems.

The data confirmed that epistemic vulnerability acts as a potent structural antecedent to both perceived civil liability exposure and fear of criminal culpability. When practitioners are confronted with opaque deep-learning models whose latent mathematical features cannot be interrogated at the point of care, their capacity to construct defensible clinical justifications is compromised. This finding validates the multi-level digital vulnerability framework articulated by Hine and Floridi (2026), demonstrating that vulnerability at the micro level (clinician lack of epistemic transparency) directly generates vulnerability at the macro level (exposure to legal sanction).

Furthermore, the statistical invariance observed between diagnostic and surgical cohorts regarding this path indicates that algorithmic opacity is a universal structural stressor. Whether interpreting cross-sectional neuroimaging or calibrating an algorithmic surgical navigation array, the inability to explain the model's underlying rationale deprives clinicians of the standard epistemic protections historically afforded by professional medical discretion (Goktas & Grzybowski, 2025; Rodrigues, 2020).

The significant negative paths between liability perceptions (civil and criminal) and protocol adoption intention underscore the structural limitations of conventional technology acceptance models in high-stakes healthcare domains. In contrast to standard consumer or enterprise environments, clinical adoption is fundamentally constrained by statutory duties of care and the omnipresent threat of malpractice litigation (Davis, 1989; Venkatesh et al., 2003).

Our findings indicate that perceived civil liability exposure operates as a primary deterrent to routine implementation, a dynamic exacerbated by the perception that prevailing tort doctrines disproportionately insulate technology developers while treating practitioners as human shock absorbers (Gourari et al., 2025; Palama, 2022).

Crucially, the confirmed partial mediation via clinical trust deficit demonstrates that liability concerns actively degrade epistemic trust in artificial intelligence systems. As clinicians recognise that institutional systems do not provide legal safe harbours, their willingness to trust algorithmic inferences declines, precipitating behavioural withdrawal from automated protocols (Khan et al., 2025; Polemi et al., 2024).

The acute impact of criminal culpability apprehension further corroborates the theoretical contentions of Mrčela and Vuletić (2025), showing that fear of gross negligence investigations or involuntary manslaughter charges linked to algorithmic failure introduces an intense chilling effect that standard performance metrics cannot counteract.

A critical contribution of this study is the empirical documentation of defensive medicine induced by algorithmic governance failures. Perceived civil liability exerted an exceptionally large positive effect on defensive utilisation behaviours, characterised by routine manual recalculations, duplicate diagnostic imaging, and exhaustive defensive charting. Rather than streamlining clinical workflows and reducing cognitive burden, unindemnified artificial intelligence deployment forces

clinicians to duplicate their labour to generate defensive paper trails (Bourne et al., 2015; Studdert et al., 2005).

This response represents a structural paradox: technologies implemented to enhance diagnostic efficiency and reduce physician fatigue instead generate operational redundancies and cognitive exhaustion as practitioners seek to insulate themselves from future litigation (Nasir et al., 2025; Pham, 2025).

The multi-group analysis revealed significant parameter divergence between diagnostic and surgical practitioners, confirming that procedural invasiveness amplifies the behavioural impact of liability exposure. While radiologists and histopathologists exhibited significant liability-induced protocol hesitation, surgeons demonstrated significantly steeper adoption resistance and higher rates of defensive clinical adaptation.

This asymmetry reflects the distinct kinetic realities of operative interventions mapped by Hashimoto et al. (2026) and Budiono et al. (2026). In surgical environments, semi-autonomous robotic actions interact directly with living tissue, where mechanical deviations, sensor drift, or miscalculated anatomical margins carry immediate risks of catastrophic patient harm or intraoperative mortality.

Because judicial frameworks continue to treat surgical robotic systems as mere instruments in the surgeon's hand, surgeons recognise that courts are unlikely to distinguish between algorithmic trajectory failure and operator negligence (Al-Zawahreh & Al-Dhanibat, 2026; Shakhathreh & Alazzam, 2025). Consequently, interventional proceduralists exhibit heightened risk aversion, frequently overriding or disengaging automated platforms to preserve personal control over operative liability.

Finally, the moderating influence of institutional resilience provides an empirical pathway toward resolving this impasse. The negative moderating effect of institutional safeguards demonstrates that when hospitals implement explicit written indemnification policies, defined clinical override boundaries, and comprehensive medicolegal training, the link between liability exposure and defensive double-checking is substantially attenuated.

These findings validate the socio-legal resilience paradigms proposed by Teo (2025) and Palama (2022), confirming that clinician vulnerability can be mitigated when healthcare institutions establish shared liability structures and clear standard-of-care frameworks that protect professional agency.

Conclusion

This study evaluated how healthcare providers perceive liability, epistemic vulnerability, and professional exposure when integrating artificial intelligence systems into diagnostic and surgical protocols. Utilizing structural equation modelling across a balanced sample of diagnostic and surgical specialists, the investigation confirmed that algorithmic opacity fuels acute perceptions of

civil and criminal liability. These legal apprehensions, in turn, degrade clinical trust, deter protocol adoption, and drive widespread defensive medical practices.

The empirical findings highlight a fundamental disconnect in modern clinical governance: while technology manufacturers and healthcare administrators advocate for rapid algorithmic adoption, the prevailing medicolegal environment continues to concentrate legal liability almost exclusively upon the individual clinician. This regulatory asymmetry transforms the clinician into an institutional liability shield, generating operational redundancies that undermine the intended clinical and economic efficiencies of artificial intelligence.

Moreover, the pronounced vulnerability observed among surgical specialists demonstrates that clinical risk scales directly with procedural invasiveness, demanding governance models tailored to the kinetic realities of interventional medicine. Ultimately, the successful and sustainable adoption of artificial intelligence in healthcare cannot rely solely on mathematical accuracy; it requires legal coherence, statutory fairness, and institutional protection that equitably distribute risk across developers, deploying facilities, and practicing clinicians.

Recommendations

Based on the empirical findings, the following structural, policy, and clinical recommendations are proposed:

1. National legislative assemblies and health law regulatory authorities must revise prevailing medical negligence and product liability frameworks to accommodate semi-autonomous decision-support systems.
2. Legal frameworks should move away from doctrines that treat algorithmic software merely as an extension of the clinician's hand.
3. Statutory regimes should implement proportionate fault allocation models that distribute liability among software developers (for training bias, data drift, and algorithm design defects), healthcare institutions (for inadequate infrastructure, maintenance, or credentialing), and clinicians (for clinical oversight).
4. Healthcare facilities and hospital management boards must establish comprehensive institutional indemnification frameworks before mandating algorithmic clinical protocols.

Declarations

Ethical Approval: This study was conducted in accordance with the Declaration of Helsinki. Formal institutional ethical approval was not required for this study as the survey was completely anonymous, voluntary, and carried no risk to participants.

Artificial Intelligence (AI) Use Statement: During the preparation of this work, the authors used Gemini to improve language, style, and clarity. Following the use of this service, the authors

thoroughly reviewed, edited, and verified the content, and take full responsibility for the final intellectual integrity of the publication.

Data Availability Statement: The datasets generated and analysed during the current study are available from the corresponding author upon reasonable request.

Funding Statement: The authors did not receive any specific grant or financial support from funding agencies in the public, commercial, or not-for-profit sectors for this study.

Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions: E. E. Aloba and I. P. Ugbor conceptualised the study and designed the research framework. E. E. Aloba conducted the core legal and ethical analysis and prepared the first draft. I. P. Ugbor oversaw the analytical methodology, validated the hypotheses, and managed project coordination as the corresponding author. M. O. Ujah contributed to the literature curation, data evaluation, and critical review of statutory frameworks. All authors contributed to the critical revision of the manuscript and approved the published version.

References

- Al-Zawahreh, M. M., & Al-Dhanibat, J. A. (2026). Civil liability of cosmetic surgeons for damages resulting from cosmetic procedures in the age of digitalization and AI-assisted surgery. In I. Younes Ibrahim Yamin & S. Reyad (Eds.), *AI-driven sustainable innovation and sustained competitive advantage* (Studies in Systems, Decision and Control, Vol. 656, pp. 29–40). Springer. https://doi.org/10.1007/978-3-032-15459-0_3
- Bourne, T., Wynants, L., Peters, M., Van Audenhove, C., Timmerman, D., Van Calster, B., & Jalmbrant, M. (2015). The impact of complaints procedures on the welfare, health and clinical practice of 7926 doctors in the UK: A cross-sectional survey. *BMJ Open*, 5(1), Article e006687. <https://doi.org/10.1136/bmjopen-2014-006687>
- Budiono, A., Mirielle, C. Y., Achmadi, A., Bangsawan, M. I., & Soekiswati, S. (2026). Rethinking patient protection and legal accountability in artificial intelligence assisted orthopedic surgery. *Journal of Human Rights, Culture and Legal System*, 6(2), 444–466. <https://doi.org/10.53955/jhcls.v6i2.995>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Elendu, C., Amaechi, D. C., Elendu, T. C., Jingwa, K. A., Okoye, O. K., John Okah, M., Ladele, J. A., Farah, A. H., & Alimi, H. A. (2023). Ethical implications of AI and robotics in

- healthcare: A review. *Medicine*, 102(50), Article e36671. <https://doi.org/10.1097/MD.00000000000036671>
- Goktas, P., & Grzybowski, A. (2025). Shaping the Future of Healthcare: Ethical Clinical Challenges and Pathways to Trustworthy AI. *Journal of Clinical Medicine*, 14(5), 1605. <https://doi.org/10.3390/jcm14051605>
- Gourari, F. M., Alazzam, F. A. F., Shakhathreh, H. J. M., Alhajeri, S. A. H., & Al Mansoori, S. A. A. (2025). Legal liability for AI-assisted radiographic diagnosis errors: A comparative review of French, Jordanian, and UAE law. *Acta Informatica Medica*, 33(3), 184–189. <https://doi.org/10.5455/aim.2025.33.184-189>
- Hashimoto, D. A., Marwaha, J. S., Lee, S. A., Schwaitzberg, S., & Duffourc, M. N. (2026). Risk and liability in the deployment of AI systems for surgery: A SAGES white paper. *Surgical Endoscopy*, 40(8), 5468–5479. <https://doi.org/10.1007/s00464-026-12881-8>
- Hine, E., & Floridi, L. (2026). Digital vulnerabilities in the age of AI: A multi-level analysis. *Philosophy & Technology*, 39(4), Article 143. <https://doi.org/10.1007/s13347-026-01150-0>
- Khan, M. M., Shah, N., Shaikh, N., Thabet, A., Alrabayah, T., & Belkhair, S. (2025). Towards secure and trusted AI in healthcare: A systematic review of emerging innovations and ethical challenges. *International Journal of Medical Informatics*, 195, Article 105780. <https://doi.org/10.1016/j.ijmedinf.2024.105780>
- Maiti, M., Kayal, P., & Vujko, A. (2025). A study on ethical implications of artificial intelligence adoption in business: Challenges and best practices. *Future Business Journal*, 11, Article 34. <https://doi.org/10.1186/s43093-025-00462-5>
- Mrčela, M., & Vuletić, I. (2025). Navigating Criminal Liability in an Era Of AI-Assisted Medicine. *Medicine, Law & Society*, 18(1), 93-108 <https://doi.org/10.18690/mls.18.1.93-108.2025>
- Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Ibrahim, S., Patil, V., Smriti, K., Shetty, S., Rai, B. P., Chlost, P., & Somani, B. K. (2022). Legal and ethical consideration in artificial intelligence in healthcare: Who takes responsibility? *Frontiers in Surgery*, 9, Article 862322. <https://doi.org/10.3389/fsurg.2022.862322>
- Nasir, M., Siddiqui, K., & Ahmed, S. (2025). Ethical-legal implications of AI-powered healthcare in critical perspective. *Frontiers in Artificial Intelligence*, 8, Article 1619463. <https://doi.org/10.3389/frai.2025.1619463>

- Palama, V. (2022). Governing High-Risk AI in Healthcare: Aligning Technical Robustness with Ethical and Legal Accountability. *International Journal of Technology, Management and Humanities*, 8(04), 65-79. <https://doi.org/10.21590/ijtmh.2022080405>
- Pham T. (2025). Ethical and legal considerations in healthcare AI: innovation and policy for safe and fair use. *Royal Society open science*, 12(5), 241873. <https://doi.org/10.1098/rsos.241873>
- Polemi, N., Praça, I., Kioskli, K., & Bécue, A. (2024). Challenges and efforts in managing AI trustworthiness risks: A state of knowledge. *Frontiers in Big Data*, 7, Article 1381163. <https://doi.org/10.3389/fdata.2024.1381163>
- Prakash, S., Balaji, J. N., Joshi, A., & Surapaneni, K. M. (2022). Ethical Conundrums in the Application of Artificial Intelligence (AI) in Healthcare—A Scoping Review of Reviews. *Journal of Personalized Medicine*, 12(11), 1914. <https://doi.org/10.3390/jpm12111914>
- Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 4, Article 100005. <https://doi.org/10.1016/j.jrt.2020.100005>
- Shakhatreh, H. J. M., & Alazzam, F. A. F. (2025). Legal frameworks for AI-driven diagnostics and robotic surgery. *Acta Informatica Medica*, 33(3), 248–252. <https://doi.org/10.5455/aim.2025.33.248-252>
- Studdert, D. M., Mello, M. M., Sage, W. M., DesRoches, C. M., Pebley, J., Zapert, K., & Brennan, T. A. (2005). Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. *JAMA*, 293(21), 2609–2617. <https://doi.org/10.1001/jama.293.21.2609>
- Teo, S. A. (2025). Artificial intelligence, human vulnerability and multi-level resilience. *Computer Law & Security Review*, 57, Article 106134. <https://doi.org/10.1016/j.clsr.2025.106134>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>