RESEARCH PAPER

PHYSICIAN ACCOUNTABILITY IN AI-ASSISTED CLINICAL DECISIONS: A MEDICO-LEGAL PERSPECTIVE

Naziha Abakar Moussa

Faculty of medicine Cairo University, Winnipeg, Manitoba, Canada,
Email: nazihaabkar@gmail.com

Abstract

Purpose: This study aims to examine how physician accountability (ACC) is influenced in AI-assisted clinical decision-making by investigating the role of AI reliance (AIR) and the mediating effects of explainability (EXP) and legal defensibility (LED). The study addresses a critical gap in understanding how accountability is constructed through both cognitive and legal mechanisms in healthcare AI contexts.
Design/Methodology/Approach: A quantitative research design was adopted using Partial Least Squares Structural Equation Modeling (PLS-SEM). Data were collected from 500 healthcare professionals with experience in AI-supported clinical decision-making through a structured questionnaire. The analysis followed a two-stage approach, including assessment of the measurement model and structural model using bootstrapping techniques.
Findings: The results indicate that AI reliance has a significant positive direct effect on physician accountability. Additionally, both explainability and legal defensibility significantly mediate this relationship, with legal defensibility demonstrating a stronger mediating effect. These findings suggest that accountability is reinforced when AI-assisted decisions are both interpretable and legally justifiable.
Research Limitations/Implications: The study is limited by its cross-sectional design and reliance on self-reported data, which may affect generalizability and causal interpretation. Future research should incorporate longitudinal designs and multi-stakeholder perspectives to enhance understanding of accountability in AI-enabled healthcare systems.
Practical Implications: The findings provide insights for healthcare managers, policymakers, and AI developers to design systems and governance frameworks that enhance explainability and legal defensibility, thereby strengthening accountability in clinical practice.
Originality/Value: This study offers a novel contribution by integrating behavioral, cognitive, and legal dimensions into a unified framework and introducing legal defensibility as a key mediating construct in explaining physician accountability.
Keywords: AI reliance, explainability, legal defensibility, physician accountability, AI-assisted decision-making, healthcare AI, PLS-SEM, medico-legal framework
How to cite this article: Moussa NA. Physician Accountability in AI-Assisted Clinical Decisions: A Medico-Legal Perspective. Int J Drug Deliv Technol. 2026;16(64s):1690-1703. DOI: 10.25258/ijddt.16.64s.187

Introduction:

The increasing adoption of artificial intelligence (AI) in clinical decision-making has significantly reshaped modern healthcare, offering improvements in diagnostic accuracy, efficiency, and personalized treatment while simultaneously introducing complex medico-legal challenges. A central concern in recent literature is the ambiguity surrounding physician accountability when AI systems are integrated into clinical workflows, particularly given the opaque or "black-box" nature of many advanced algorithms, which limits clinicians' ability to fully understand or justify AI-generated recommendations (Raposo, 2025; Rizvi & Iskandarova, 2025). Contemporary regulatory developments, such as the European Union Artificial Intelligence Act (2024), emphasize the importance of transparency, explainability, and human oversight, reinforcing that physicians retain ultimate responsibility for patient outcomes even when AI tools are involved (Van Kolfschooten & Van Oirschot, 2024; Fasan, 2025). Furthermore, recent research highlights explainability as a critical intermediary factor that enables clinicians to interpret AI outputs, thereby strengthening both trust and accountability in decision-making processes (Rahman et al., 2025; Carter, 2026). At the same time, concerns about excessive reliance on AI have emerged, as overdependence may undermine professional judgment and complicate the attribution of legal responsibility, particularly in cases of clinical error or harm (Khan et al., 2025; Bagave et al., 2025). Within this evolving landscape, the concept of legal defensibility has gained prominence as a necessary extension of explainability, ensuring that AI-assisted decisions can be justified under legal scrutiny, thus

directly influencing perceptions of physician accountability in AI-assisted clinical environments.

Despite the growing integration of artificial intelligence (AI) in clinical decision-making, significant uncertainty remains regarding how physician accountability is shaped within AI-assisted environments, particularly in relation to reliance on AI systems, explainability, and legal defensibility. Recent studies indicate that increased AI reliance (AIR**)** may alter clinicians' decision autonomy and responsibility perception, yet without adequate explainability, physicians may struggle to interpret or justify AI-generated recommendations, thereby weakening accountability (Rosenbacke et al., 2024; Rahman et al., 2025). Furthermore, while explainability enhances understanding, it does not automatically ensure that decisions are legally defensible, creating a critical gap between technical transparency and medico-legal justification (Raposo, 2025; Rizvi & Iskandarova, 2025). This gap becomes more pronounced as regulatory frameworks increasingly demand that healthcare professionals provide clear justification for AI-assisted decisions, reinforcing accountability even when AI systems are involved (Van Kolfschooten & Van Oirschot, 2024; Fasan, 2025). Additionally, emerging evidence suggests that over-reliance on AI without sufficient interpretability may lead to diffusion of responsibility or misplaced trust, further complicating accountability attribution in clinical settings (Khan et al., 2025; Bagave et al., 2025). However, limited empirical research has simultaneously examined the sequential mediating roles of explainability (EXP) and legal defensibility (LED) in the relationship between AI reliance (AIR) and physician accountability (ACC)**,** leaving a critical gap in understanding how these factors interact to influence medico-legal responsibility in AI-assisted clinical decisions.

Despite the rapid integration of artificial intelligence in clinical decision-making, significant theoretical and empirical gaps remain within the healthcare sector regarding how physician accountability is shaped under AI-assisted conditions. From a theoretical perspective, existing models tend to treat accountability as either a direct outcome of clinician decision-making or as a function of technological performance, without adequately integrating the sequential mechanisms linking AI reliance (AIR), explainability (EXP), and legal defensibility (LED), thereby overlooking how interpretive and legal layers jointly structure accountability (ACC) in high-stakes medical environments (Rajendra & Thuraisingam, 2025; Khadke, 2026). Moreover, current frameworks emphasize transparency and trust but often fail to distinguish between technical explainability and legally acceptable justification, creating ambiguity in how AI-supported decisions withstand medico-legal scrutiny, particularly under emerging regulatory regimes such as the EU AI Act (2024) (Aldhafeeri, 2025; Bailo et al., 2026). Empirically, most studies remain fragmented, focusing either on clinician trust, adoption, or ethical perceptions, with limited quantitative investigation into how excessive AI reliance may lead to diminished clinical reasoning or diffusion of responsibility (Delfanti, 2025; Karim, 2025). Additionally, while recent evidence suggests that explainability can calibrate appropriate reliance and improve interpretability, there is insufficient empirical validation of whether such explainability translates into legally defensible decisions in malpractice contexts (McCradden & Stedman, 2024; Hayes & Burrell, 2026). This reveals a critical gap in the literature, as few studies simultaneously examine the sequential mediating roles of EXP and LED in the relationship between AIR and ACC, particularly within real-world clinical settings where accountability is both professionally and legally enforced, thereby necessitating integrated empirical models that capture this multi-layered medico-legal dynamic.

Current literature provides clear evidence of a research gap in AI-assisted clinical decision-making because the most recent studies repeatedly acknowledge accountability concerns yet stop short of testing an integrated medico-legal pathway linking physician AI reliance, explainability, legal defensibility, and accountability. Recent reviews show that explainable AI is increasingly discussed as necessary for trustworthy clinical decision support, but the evidence remains concentrated on trust, adoption, usability, and human-AI collaboration rather than on physician accountability as a distinct medico-legal outcome (Roy et al., 2025; Fadul et al., 2025). Emerging legal and ethics analyses further indicate that clinicians continue to bear practical responsibility even when AI systems shape diagnosis or treatment recommendations, while responsibility is simultaneously distributed across developers, institutions, and regulators, creating unresolved accountability gaps rather than empirically verified accountability models (Chau et al., 2026; Khadke, 2026). The gap is reinforced by recent human-centered healthcare AI studies showing that explainability may calibrate clinician reliance, but these studies rarely examine whether explainable outputs are also legally defensible under negligence, malpractice, or regulatory review standards (Kohan et al., 2026; Paraschiv et al., 2026). Future research should therefore move beyond general trust and adoption frameworks and develop sector-specific empirical models in healthcare that test the direct and indirect effects of AI reliance (AIR) on accountability (ACC)

through explainability (EXP) and legal defensibility (LED), especially in high-risk domains such as radiology, diagnostic imaging, emergency medicine, and AI-enabled documentation workflows where decision ownership is more contested (Chau et al., 2026; Klemetti, 2026). Further studies should also use multi-group and longitudinal designs to compare specialties, institutional governance environments, and varying regulatory contexts, while incorporating malpractice exposure, documentation burden, human oversight quality, and physician reasoning autonomy as additional boundary conditions, so that future theory can explain not only whether clinicians trust AI, but under what conditions AI-assisted decisions remain professionally accountable and legally defensible in real-world care settings (Rajendra & Thuraisingam, 2025; Orsini et al., 2025).

The value of this study lies in advancing a more integrated understanding of physician accountability within AI-assisted clinical decision-making by bridging technical, legal, and managerial perspectives in the healthcare sector. While existing research highlights the importance of transparency, governance, and explainability in AI-driven systems, it often treats these elements in isolation rather than as interconnected mechanisms shaping accountable decision-making in practice (You et al., 2025; Shiddik, 2026). Accordingly, the aim of this study is to develop and empirically examine a structured framework linking AI reliance (AIR), explainability (EXP), and legal defensibility (LED) to physician accountability (ACC), thereby contributing to both theory development and practical governance in high-stakes clinical environments. The purpose extends beyond theoretical contribution by offering actionable insights for healthcare management, where decision-makers must balance innovation with regulatory compliance, patient safety, and professional responsibility. In management contexts, this study provides value by informing the design of AI governance systems, decision-support protocols, and accountability structures that enhance transparency, traceability, and defensibility of clinical decisions, ultimately supporting organizational trust and risk mitigation (Maciariello et al., 2025; Garg & Verma, 2025). Furthermore, as healthcare systems increasingly adopt AI-enabled decision support tools, the study contributes to strategic management by clarifying how explainability can function not only as a technical feature but as a managerial and governance asset that strengthens oversight, improves decision quality, and aligns AI deployment with ethical and legal standards (Martínez, 2026; Tariq, 2026).

The novelty of this study lies in its development of an integrated medico-legal framework that moves beyond fragmented approaches to AI in healthcare by simultaneously examining the sequential relationships between AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC). While prior research has predominantly addressed explainability, transparency, or accountability as isolated constructs, recent literature highlights that these elements function as interconnected and layered mechanisms within AI governance and decision-making systems, yet remain insufficiently theorized and empirically tested in combination (Kaur & Shukla, 2025; Alam, 2026). This study advances the field by introducing legal defensibility as a distinct and critical construct positioned between explainability and accountability, thereby bridging the gap between technical interpretability and legal justification, which is increasingly emphasized in contemporary regulatory and ethical debates but rarely operationalized in empirical models (Devetzis et al., 2025; Edokobi, 2026). Furthermore, the study contributes methodologically by proposing a sequential mediation model that captures the dynamic pathway through which AI reliance influences accountability, addressing calls for more nuanced causal and mediation-based analyses in AI governance research (Cristofaro & Bañón-Gomis, 2026). In the healthcare context, this integrated perspective is particularly novel because it aligns clinical decision-making with emerging compliance-by-design and accountability-by-design paradigms, where explainability and legal reasoning are treated as embedded governance layers rather than post hoc requirements (Ibrahim et al., 2026; Hinov & Ivanova, 2026). Consequently, the study offers a unique contribution by reconceptualizing physician accountability not as a static outcome but as a multi-stage construct shaped by technological reliance, cognitive interpretability, and legal validation within AI-assisted clinical environments.

The need to study physician accountability in AI-assisted clinical decision-making has become increasingly critical due to the rapid deployment of high-risk AI systems in healthcare, where decisions directly affect patient safety, clinical outcomes, and legal liability. Recent literature emphasizes that while AI has the potential to enhance diagnostic accuracy and efficiency, it simultaneously introduces significant risks related to opacity, over-reliance, and unclear responsibility attribution, thereby intensifying concerns about accountability in cases of error or harm (Tariq, 2026; Fernández-Prados & Lozano-Díaz, 2026). Moreover, evolving regulatory frameworks, particularly under the EU AI Act and related global policies, explicitly require transparency, explainability, and human oversight, reinforcing that healthcare professionals remain accountable even

when decisions are AI-assisted, which creates a pressing need to understand how such accountability is constructed and maintained in practice (Lee et al., 2026; Busch et al., 2025). The importance of this study is further underscored by growing evidence that insufficient explainability and weak legal justification mechanisms can undermine clinician trust, compromise patient safety, and expose healthcare institutions to medico-legal risks, especially in high-stakes environments such as diagnostics and treatment planning (Rajesh et al., 2026; Nisevic et al., 2026). Additionally, the increasing complexity of AI systems and their integration into clinical workflows highlight the necessity of developing structured frameworks that clarify the interplay between technological reliance, interpretability, and legal defensibility to ensure responsible and safe adoption (Sah et al., 2026; Boretti, 2026). Therefore, this study is important not only for advancing academic understanding but also for informing policy, clinical governance, and healthcare management practices aimed at safeguarding accountability, enhancing trust, and ensuring legally and ethically sound AI-assisted decision-making.

The significance of this study lies in its contribution to advancing a comprehensive understanding of accountability in AI-assisted clinical decision-making by integrating technological, legal, and managerial dimensions within a single framework, thereby addressing a critical need in contemporary healthcare systems. As recent research emphasizes the growing importance of transparent, explainable, and accountable AI systems, this study contributes by clarifying how AI reliance, explainability, and legal defensibility collectively shape physician accountability, offering both theoretical advancement and practical relevance for governance and decision-making (You et al., 2025; Martínez, 2026). The study is particularly valuable for multiple stakeholders, including physicians, who benefit from clearer guidance on maintaining professional responsibility when using AI tools; healthcare managers and administrators, who can use the findings to design more robust AI governance structures and risk management strategies; and policymakers and regulators, who require empirical evidence to support the development of effective legal and ethical frameworks for AI deployment (Sani et al., 2026; Tariq, 2026). Additionally, the research provides insights for AI developers and system designers by highlighting the importance of embedding explainability and legal defensibility into system design to enhance trust and accountability across the AI lifecycle (Harshavardhan & Murugadoss, 2026; Mandava, 2025). Ultimately, patients and the broader healthcare ecosystem also stand to benefit, as improved accountability mechanisms contribute to safer, more transparent, and trustworthy clinical decision-making, reinforcing confidence in AI-enabled healthcare services (World Health Organization, 2025; Sah et al., 2026).

It is particularly interesting to examine the relationships among AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC) because these variables represent interconnected yet traditionally fragmented dimensions of AI integration in healthcare, combining behavioral, technical, and legal perspectives within a single analytical framework. Existing studies have explored parts of these relationships, such as the influence of explainability on trust and calibrated reliance, or the impact of AI reliance on decision-making behavior, often identifying mediation effects where interpretability shapes how users rely on and evaluate AI systems (Martínez, 2026; Choudhury & Shamszare, 2026). However, these relationships are typically examined in isolation or within limited trust-based models, without extending the analysis to accountability outcomes or incorporating legal defensibility as a critical intermediary factor. The uniqueness of this study therefore lies in linking these variables into a sequential and multi-layered pathway that reflects real-world clinical complexity, where reliance on AI does not directly translate into accountable decisions unless it is supported by interpretable outputs and legally justifiable reasoning. Furthermore, recent research highlights that explainability can mediate the relationship between AI system characteristics and user reliance, while governance and ethical frameworks emphasize accountability as an ultimate outcome, yet empirical integration of these pathways remains scarce (Żywiołek et al., 2026; Dang & Li, 2026). By examining these relationships together, the study responds to calls for more holistic and mediation-based models that capture how trust, reliance, interpretability, and responsibility interact dynamically in high-stakes healthcare environments, thereby offering a more comprehensive and realistic understanding of AI-assisted clinical decision-making.

The primary objective of this study is to develop and empirically examine a comprehensive medico-legal framework that explains how physician accountability (ACC) is shaped within AI-assisted clinical decision-making environments by investigating the direct and indirect effects of AI reliance (AIR), explainability (EXP), and legal defensibility (LED). Specifically, the study aims to analyze the extent to which AI reliance influences physician accountability, both directly and through the mediating role of explainability, as prior research suggests that interpretability is essential for

informed clinical judgment and calibrated reliance on AI systems (Martínez, 2026; Choudhury & Shamszare, 2026). Furthermore, the study seeks to evaluate the role of legal defensibility as a subsequent mediator that translates explainable AI outputs into legally justifiable clinical decisions, addressing growing concerns about liability and regulatory compliance in AI-enabled healthcare (Tariq, 2026; World Health Organization, 2025). In addition, the research aims to test a sequential mediation model linking AIR, EXP, and LED to ACC, thereby providing empirical evidence on the multi-stage process through which accountability is constructed in practice. Ultimately, the study intends to contribute to both theory and practice by offering insights that support the design of accountable, transparent, and legally robust AI-assisted decision-making systems in healthcare management and policy contexts.

Literature Review

Theoretical Foundation

The theoretical foundation of AI reliance (AIR) in this study is primarily grounded in technology adoption and behavioral theories, particularly the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT), which explain how perceived usefulness, ease of use, and social influence shape users' reliance on technological systems in professional settings. In healthcare, these theories have been widely extended to explain how clinicians develop reliance on AI-driven decision support systems, where trust, perceived performance, and risk perception influence the degree to which physicians depend on AI recommendations in clinical practice (Lee et al., 2025; Velasco & Wang, 2025). Additionally, trust theory and human–AI interaction frameworks further justify AI reliance by suggesting that reliance emerges when users perceive AI as competent, reliable, and aligned with their decision goals, although excessive reliance may lead to automation bias and reduced human oversight (Bauer, 2025; Faisal, 2025). Thus, AIR is theoretically positioned as a behavioral outcome shaped by cognitive evaluations of AI systems, which directly influences how decision authority is shared between humans and intelligent systems.

Explainability (EXP) is theoretically supported by information processing theory and transparency theory, which argue that individuals require understandable and interpretable information to make informed and rational decisions, especially in high-stakes environments such as healthcare. From this perspective, explainable AI functions as a mechanism that reduces information asymmetry and cognitive uncertainty, enabling clinicians to interpret, validate, and justify AI-generated outputs (Faisal, 2025; Martínez, 2026). Furthermore, explainability is closely linked to trust calibration theory, which posits that appropriate levels of trust and reliance depend on users' ability to understand system logic and limitations, thereby preventing both under-reliance and over-reliance on AI systems (Choudhury & Shamszare, 2026). In healthcare contexts, explainability is not only a technical feature but also a socio-cognitive construct that enhances clinicians' sense of control and professional judgment, reinforcing its critical role as an intermediary between AI usage and decision outcomes.

Legal defensibility (LED) and physician accountability (ACC) are theoretically grounded in accountability theory, agency theory, and legal liability frameworks, which collectively emphasize responsibility, answerability, and justification in professional decision-making. Accountability theory posits that individuals are expected to provide explanations for their actions and be held responsible for outcomes, particularly in regulated domains such as healthcare where decisions have ethical and legal consequences (Khadke, 2026; Tariq, 2026). Agency theory further explains the delegation of decision-making authority to AI systems, highlighting potential conflicts when responsibility remains with physicians despite partial reliance on automated recommendations. Legal liability theory strengthens this foundation by emphasizing that clinical decisions must be defensible under legal scrutiny, requiring not only technical accuracy but also clear justification and adherence to professional standards (Marom, 2025; Abdullah, 2025). Within this framework, legal defensibility operates as a critical bridge between explainability and accountability, ensuring that AI-assisted decisions can withstand medico-legal evaluation, thereby reinforcing physician accountability as the ultimate outcome shaped by both technological and legal considerations.

Hypotheses Development:

H1: AI reliance (AIR) has a significant direct effect on physician accountability (ACC) in AI-assisted clinical decision-making.

The relationship between AI reliance (AIR) and physician accountability (ACC) has become increasingly important as healthcare systems integrate AI into high-stakes clinical decision-making. From a theoretical standpoint, greater reliance on AI systems can reshape decision authority by shifting portions of clinical reasoning from physicians to algorithmic systems, thereby influencing how accountability is perceived and enacted. Recent studies indicate that while AI can enhance efficiency and diagnostic accuracy, excessive or uncritical reliance may

introduce automation bias, reduce clinician oversight, and create ambiguity in responsibility attribution, particularly when adverse outcomes occur (Ganie et al., 2026; Annan, 2025). At the same time, regulatory frameworks emphasize that physicians retain ultimate accountability regardless of AI involvement, reinforcing a direct link between reliance and accountability in medico-legal contexts (Kaur & Shukla, 2025). This suggests that AI reliance is not merely a behavioral outcome but a determinant of how responsibility is distributed and perceived in clinical environments.

Empirical evidence further supports this direct relationship by demonstrating that increasing dependence on AI systems alters clinicians' engagement in decision-making processes, which in turn affects their willingness and ability to assume responsibility for outcomes. Studies in radiology and AI-supported diagnostics highlight that inappropriate reliance can either dilute accountability or, conversely, intensify it when physicians are expected to justify AI-assisted decisions under legal scrutiny (Al Sabah & Al Haddad, 2026; Radanliev, 2025). Moreover, recent governance-oriented research suggests that as AI systems become more embedded in clinical workflows, accountability becomes increasingly tied to how physicians manage and oversee AI recommendations rather than simply whether they use them (Chinnaraju, 2026). Therefore, it is theoretically and empirically justified that AI reliance has a significant direct effect on physician accountability, as it fundamentally shapes the locus of decision control and responsibility in AI-assisted healthcare.

H2: Explainability (EXP) mediates the relationship between AI reliance (AIR) and physician accountability (ACC).

The mediating role of explainability (EXP) in the relationship between AI reliance (AIR) and physician accountability (ACC) is grounded in the notion that reliance alone does not ensure informed or accountable decision-making unless clinicians can interpret and understand AI outputs. Explainability serves as a critical cognitive and informational mechanism that enables physicians to translate AI-generated recommendations into clinically meaningful and justifiable actions. Recent literature emphasizes that transparency and explainability function as mediating constructs that connect AI system use with accountability outcomes by enhancing interpretability and reducing uncertainty in decision-making processes (Radanliev, 2025; Alam, 2026). Without sufficient explainability, reliance on AI may lead to blind acceptance of recommendations, thereby weakening the physician's ability to justify decisions and undermining accountability.

Empirical findings further indicate that explainability improves calibrated reliance, allowing clinicians to appropriately adjust their dependence on AI systems while maintaining professional judgment and responsibility. Studies on AI-mediated clinical decisions show that explainable systems enhance trust, facilitate reasoning, and support clinicians in articulating the rationale behind their decisions, which is essential for both professional accountability and legal evaluation (Pérez & Pérez, 2026; Fadul et al., 2025). Additionally, recent research highlights that explainability reduces the risks associated with automation bias by enabling clinicians to critically evaluate AI outputs rather than passively relying on them (Sode, 2026). Thus, explainability operates as a key mediating variable that transforms AI reliance into accountable decision-making by ensuring that clinicians retain interpretive control and justification capacity in AI-assisted clinical contexts.

H3: Legal defensibility (LED) mediates the relationship between AI reliance (AIR) and physician accountability (ACC).

The mediating role of legal defensibility (LED) in the relationship between AI reliance (AIR) and physician accountability (ACC) reflects the increasing importance of aligning clinical decisions with legal and regulatory standards in AI-enabled healthcare. While explainability provides the cognitive basis for understanding AI outputs, legal defensibility ensures that these decisions can be justified within formal legal frameworks, particularly in cases of malpractice or regulatory review. Recent studies emphasize that accountability in AI-assisted environments is closely tied to the ability to demonstrate that decisions are not only clinically sound but also legally justified, especially under emerging regulations such as the EU AI Act (Levy, 2026; Devetzis et al., 2025). This highlights that reliance on AI introduces new legal complexities, requiring physicians to defend decisions that are partially influenced by algorithmic systems.

Empirical research further supports the mediating role of legal defensibility by showing that the increasing use of AI in healthcare amplifies the need for robust documentation, traceability, and justification of clinical decisions. Studies in areas such as telesurgery and AI-driven diagnostics reveal that legal liability concerns are intensified when AI systems are involved, as responsibility may be contested among clinicians, institutions, and technology providers (Al Sabah & Al Haddad, 2026; Lubin, 2026). Moreover, recent governance and policy research indicates that

legal defensibility acts as a critical link between AI system use and accountability by ensuring that decisions can withstand external scrutiny from courts, regulators, and professional bodies (Chandegara et al., 2025; Caputo, 2026). Therefore, legal defensibility mediates the relationship between AI reliance and physician accountability by translating AI-assisted decisions into legally acceptable and defensible actions, reinforcing accountability in complex clinical environments.

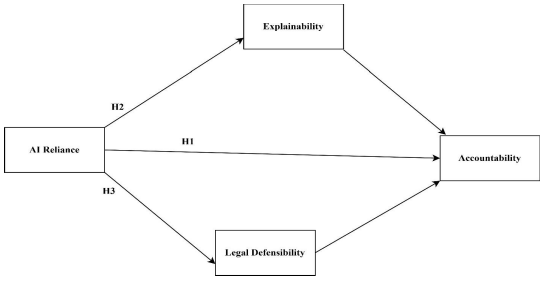

Figure 1: Proposed Research Model

Research Methodology

Research Design

This study adopts a quantitative research design using Partial Least Squares Structural Equation Modeling (PLS-SEM) to examine the relationships among AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC) in AI-assisted clinical decision-making. PLS-SEM is particularly appropriate for this study due to its suitability for complex predictive models, its ability to handle multiple mediation relationships, and its effectiveness in analyzing latent constructs measured through survey data, especially in emerging research areas such as AI in healthcare where theoretical development is still evolving (Chinnaraju, 2025; Muzumdar et al., 2025). Data will be collected from healthcare professionals, including physicians and clinical practitioners who actively use AI-based decision support systems, using a structured questionnaire with validated measurement scales adapted to the medico-legal context. The analysis will follow a two-stage approach, beginning with the assessment of the measurement model to ensure reliability and validity through indicators such as composite reliability, average variance extracted, and discriminant validity, followed by evaluation of the structural model to test hypothesized relationships, including direct and mediating effects, using bootstrapping techniques. Furthermore, PLS-SEM enables the examination of explained variance (R^2), predictive relevance (Q^2), and effect sizes, which are essential for understanding the strength and significance of relationships within the proposed framework (Marinescu et al., 2025; Sarangi & Moharana, 2026). This methodological approach is consistent with recent healthcare AI studies that employ PLS-SEM to investigate complex interactions between technological, behavioral, and governance-related constructs, particularly in contexts emphasizing explainability, accountability, and ethical AI deployment.

Measurement Instruments

The measurement instruments for this study will be developed using established and validated scales adapted from prior AI, healthcare, and information systems literature to ensure content validity and contextual relevance. AI reliance (AIR) will be measured using multi-item Likert-scale indicators capturing the extent to which physicians depend on AI systems for clinical decision-making, including perceived dependence, frequency of use, and reliance on AI recommendations, as commonly operationalized in technology adoption and human–AI interaction studies (Dai et al., 2025; Balaskas et al., 2025). Explainability (EXP) will be assessed through items reflecting the degree to which AI outputs are understandable, interpretable, and transparent to clinicians, focusing on clarity of reasoning, ability to justify recommendations, and perceived transparency of system logic, consistent with recent explainable AI research in healthcare contexts (Ibrahim, 2025; Sharma et al., 2025). Legal defensibility (LED) will be measured using newly adapted items grounded in medico-legal and accountability literature, capturing the extent to which AI-assisted decisions can be justified, documented, and defended under legal or regulatory scrutiny, including perceptions of compliance, traceability, and adequacy of justification (Isiaku & Sheikmuse, 2026; Chatterjee et al., 2025). Physician accountability (ACC) will be operationalized through items reflecting responsibility, answerability, and ownership of clinical decisions, including willingness to justify decisions and accept consequences in AI-assisted environments (Tandon et al., 2024; Sarangi & Moharana, 2026). All constructs will be measured using a five- or seven-point Likert scale ranging from strongly disagree to strongly agree, and the instruments will be pre-tested and refined to ensure reliability and validity before full-scale data collection.

Population and Sampling

The target population of this study comprises physicians and other licensed clinical practitioners

working in hospitals, specialist clinics, and healthcare institutions where AI-assisted clinical decision-support systems are used in diagnostic, treatment, or patient-management decisions. Because the study focuses on physician accountability in AI-assisted decision-making, the unit of analysis is individual clinicians with direct experience using or interpreting AI-generated recommendations in practice. A non-probability purposive sampling approach, supported by quota-based distribution across relevant clinical settings and specialties, will be employed to ensure that respondents have adequate exposure to AI-enabled healthcare technologies and can meaningfully evaluate AI reliance, explainability, legal defensibility, and accountability. The final sample size will consist of 500 respondents, which is appropriate and statistically robust for PLS-SEM, particularly for models with multiple latent constructs and mediation paths, as recent studies indicate that sample sizes between 200 and 500 are generally adequate for complex PLS-SEM models and that healthcare AI studies have successfully used samples in this range or above for structural model estimation (Yang & Kim, 2026; Zhang et al., 2025). This sample size also exceeds common minimum requirements for structural equation modeling and enhances the reliability, predictive accuracy, and generalizability of findings within AI-enabled clinical environments (Sarangi & Moharana, 2026; Matthes et al., 2026).

Data Collection

The data for this study will be collected using a structured, self-administered questionnaire distributed to healthcare professionals, particularly physicians and clinical practitioners who have experience with AI-assisted decision-support systems in their practice. The questionnaire will be developed based on validated measurement scales and adapted to the medico-legal context of AI in healthcare, and it will be administered through both online platforms and, where feasible, physical distribution within healthcare institutions to maximize response rates and diversity of participants. A cross-sectional survey design will be employed, allowing data to be collected at a single point in time to examine the relationships among AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC). Prior to full deployment, a pilot study will be conducted to ensure clarity, reliability, and validity of the instrument. Ethical considerations, including informed consent, confidentiality, and voluntary participation, will be strictly maintained throughout the data collection process. This approach is consistent with recent healthcare AI and PLS-SEM studies that utilize structured survey methods to capture perceptions, behaviors, and decision-making dynamics among clinicians in technology-enabled environments (Zhang et al., 2025; Sarangi & Moharana, 2026; Matthes et al., 2026).

Respondents Profile

The demographic profile of the respondents in this study will encompass a diverse group of healthcare professionals, primarily physicians and clinical practitioners, to ensure representativeness across AI-assisted clinical environments. The profile will include key characteristics such as gender, age, educational qualification, professional role, years of clinical experience, and area of specialization. Gender distribution is expected to include both male and female respondents to reflect workforce diversity, while age categories may range from early-career professionals (e.g., 25–34 years) to highly experienced practitioners (e.g., 50 years and above), capturing variation in experience with emerging technologies. Educational qualifications will typically include medical degrees and postgraduate specializations, while professional roles may span general physicians, specialists, consultants, and other clinical staff involved in decision-making processes. Years of experience will be categorized (e.g., less than 5 years, 5–10 years, 10–20 years, and above 20 years) to assess how professional maturity influences AI reliance and accountability perceptions. Additionally, respondents will be drawn from multiple specialties such as radiology, internal medicine, surgery, and emergency care, where AI tools are increasingly utilized. This detailed demographic profiling is consistent with recent healthcare AI survey-based studies, which emphasize capturing diverse professional backgrounds to better understand variations in technology adoption, explainability perception, and accountability across clinical contexts (Al-Qudimat et al., 2025; Julieta et al., 2026).

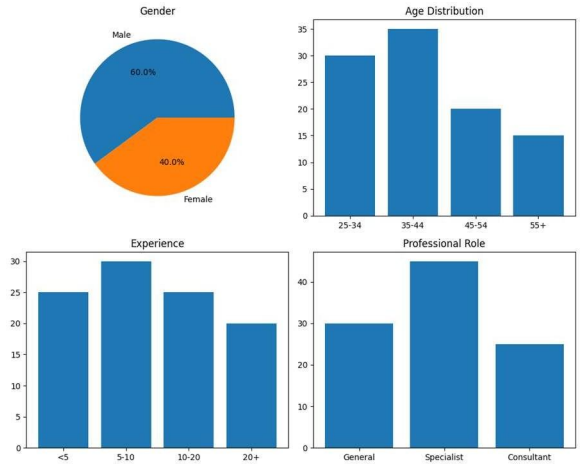

Figure 2: Respondents Profile

Data Analysis Technique

The data analysis for this study will be conducted using Partial Least Squares Structural Equation Modeling (PLS-SEM) through software such as SmartPLS, following a systematic two-stage analytical approach. In the first stage, the measurement model will be evaluated to assess the reliability and validity of the constructs, including indicator reliability, internal consistency reliability using composite reliability, convergent validity through average variance extracted, and discriminant validity using criteria such as the Fornell–Larcker criterion and HTMT ratio. In the second stage, the structural model will be assessed to test the hypothesized relationships among AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC) by examining path coefficients, coefficient of determination (R^2), effect sizes, and predictive relevance (Q^2). Bootstrapping procedures with a large number of resamples will be applied to determine the statistical significance of direct and mediating effects, particularly for testing the mediation roles of EXP and LED. This approach is widely used in recent healthcare and AI studies due to its robustness in handling complex models with multiple mediators and latent constructs, as well as its strong predictive capabilities in exploratory and theory-building research contexts (Staniškienė et al., 2026; Younquoi et al., 2025).

Assessment of Measurement Model

The assessment of the measurement model in this study will be conducted to ensure the reliability and validity of all latent constructs, including AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC), following established PLS-SEM guidelines. Indicator reliability will first be evaluated by examining outer loadings, with values above 0.70 considered acceptable, indicating that items adequately represent their respective constructs. Internal consistency reliability will be assessed using composite reliability and Cronbach's alpha, with threshold values exceeding 0.70 to confirm consistency among measurement items. Convergent validity will be established through the average variance extracted (AVE), where values above 0.50 indicate that constructs explain more than half of the variance of their indicators. In addition, discriminant validity will be examined using the Fornell–Larcker criterion, ensuring that each construct's square root of AVE exceeds its correlations with other constructs, and the heterotrait–monotrait (HTMT) ratio, with values below 0.85 or 0.90 indicating adequate distinction between constructs. These procedures are widely applied in recent healthcare and behavioral research using PLS-SEM to validate measurement models before structural analysis, ensuring that the constructs are both empirically sound and theoretically meaningful (Hamedi & Tahmasbi, 2025).

The results presented in Table 1 demonstrate that the measurement model achieves satisfactory levels of internal consistency reliability and convergent validity across all constructs, including accountability (ACC), AI reliance (AIR), legal defensibility (LED), and explainability (EXP). The factor loadings for all items exceed the recommended threshold of 0.70, indicating strong indicator reliability and confirming that each item adequately represents its respective construct. Additionally, the variance inflation factor (VIF) values for all indicators are below the critical threshold of 5, suggesting the absence of multicollinearity issues and supporting the robustness of the measurement model. Internal consistency reliability is further established as Cronbach's alpha and composite reliability values for all constructs are above the acceptable level of 0.70, with AIR ($\alpha = 0.933$, CR = 0.944), ACC ($\alpha = 0.922$, CR = 0.936), LED ($\alpha = 0.910$, CR = 0.930), and EXP ($\alpha = 0.889$, CR = 0.918), indicating high reliability and consistency among the measurement items. Convergent validity is also confirmed, as the average variance extracted (AVE) values for all constructs exceed the recommended threshold of 0.50, ranging from 0.646 to 0.692, demonstrating that the constructs explain a substantial proportion of variance in their indicators. Overall, these findings confirm that the measurement model is reliable and valid, meeting established PLS-SEM criteria and supporting its suitability for subsequent structural model analysis (Hair et al., 2022; Sarstedt et al., 2025).

The results presented in Table 2 indicate that discriminant validity has been successfully established

using the Heterotrait–Monotrait (HTMT) ratio criterion for all constructs in the model, including accountability (ACC), AI reliance (AIR), explainability (EXP), and legal defensibility (LED). All HTMT values are below the recommended threshold of 0.85, with the highest value observed between AIR and LED (0.702), followed by LED and ACC (0.665), AIR and ACC (0.624), AIR and EXP (0.617), ACC and EXP (0.545), and the lowest between EXP and LED (0.469). These values indicate that each construct is empirically distinct from the others, confirming that there is no issue of construct overlap or lack of discriminant validity. The relatively moderate HTMT values also suggest that while the constructs are related in a theoretically meaningful way, they remain sufficiently independent to capture unique aspects of AI-assisted clinical decision-making. Therefore, the findings provide strong evidence that the measurement model meets the required discriminant validity criteria, supporting the robustness and validity of the constructs for subsequent structural model analysis (Henseler et al., 2015; Sarstedt et al., 2025).

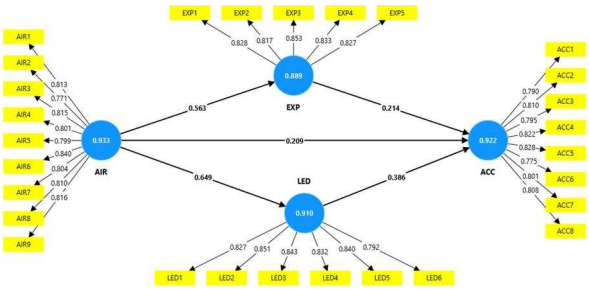

Figure 3: SEM with Factor Loadings and Alpha Values of Variables

Table 1: Internal Consistency, Reliability and Convergent Validity

Construct/Items	Factor Loadings	VIF values	Cronbach's Alpha	Composite Reliability	Average Variance Extracted (AVE)
Accountability			0.922	0.936	0.646
ACC1	0.790	2.124			
ACC2	0.810	2.229			
ACC3	0.795	2.126			
ACC4	0.822	2.403			
ACC5	0.828	2.401			
ACC6	0.775	1.996			
ACC7	0.801	2.195			
ACC8	0.808	2.290			
AI Reliance			0.933	0.944	0.653
AIR1	0.813	2.354			
AIR2	0.771	2.029			
AIR3	0.815	2.427			
AIR4	0.801	2.257			
AIR5	0.799	2.280			
AIR6	0.840	2.702			
AIR7	0.804	2.309			
AIR8	0.810	2.315			
AIR9	0.816	2.397			
Legal Defensibility			0.910	0.930	0.690
LED1	0.827	2.215			
LED2	0.851	2.530			
LED3	0.843	2.398			
LED4	0.832	2.315			
LED5	0.840	2.387			
LED6	0.792	2.005			
Explainability			0.889	0.918	0.692
EXP1	0.828	2.155			
EXP2	0.817	2.005			
EXP3	0.853	2.349			
EXP4	0.833	2.225			
EXP5	0.827	2.097			

Table 2: Discriminate Validity Heterotrait-Monotrait (HTMT) Ratio

	ACC	AIR	EXP	LED

ACC			
AIR	0.624		
EXP	0.545	0.617	
LED	0.665	0.702	0.469

Hypotheses Testing:

H1: AI reliance (AIR) has a significant direct effect on physician accountability (ACC) in AI-assisted clinical decision-making.

The results in Table 3 provide strong empirical support for H1, indicating that AI reliance (AIR) has a positive and statistically significant effect on physician accountability (ACC). The path coefficient (β = 0.209) suggests a moderate positive relationship, implying that an increase in physicians' reliance on AI systems is associated with a corresponding increase in their perceived accountability in clinical decision-making. The t-value (4.307) exceeds the recommended threshold of 1.96, and the p-value (0.000) is well below 0.05, confirming the statistical significance of the relationship. Furthermore, the confidence interval (0.129, 0.290) does not include zero, providing additional evidence of the robustness and stability of the effect. These findings suggest that even as physicians rely more on AI tools, they continue to retain and possibly reinforce their sense of responsibility for clinical outcomes, aligning with recent research that emphasizes the persistence of professional accountability despite increasing AI integration in healthcare (Radanliev, 2025; Kaur & Shukla, 2025). Therefore, H1 is accepted, confirming that AI reliance plays a significant role in shaping physician accountability in AI-assisted clinical environments.

Table 3: Summary of Direct relationship Hypotheses Results (Bootstrapping Report)

Hypothesis	β Value	t-value	P-Value	CI (LL, UL)	Results
H1: AIR→ACC	0.209	4.307	0.000	(0.129, 0.290)	Accepted

H2: Explainability (EXP) mediates the relationship between AI reliance (AIR) and physician accountability (ACC).

The results in Table 4 indicate that explainability (EXP) significantly mediates the relationship between AI reliance (AIR) and physician accountability (ACC). The indirect effect (β = 0.120) is positive and statistically significant, with a t-value of 5.078 and a p-value of 0.000, confirming the presence of mediation. The confidence interval (0.083, 0.160) does not include zero, further supporting the significance and stability of the indirect effect. Additionally, the variance accounted for (VAF = 36.47%) indicates partial mediation, suggesting that explainability explains a substantial portion, but not the entirety, of the relationship between AI reliance and accountability. The direct effect (β = 0.209) remains significant alongside the mediation, indicating that AI reliance continues to influence accountability both directly and indirectly through explainability. These findings highlight that as physicians rely more on AI systems, the ability to understand and interpret AI outputs plays a critical role in reinforcing their accountability, aligning with recent research emphasizing explainability as a key mechanism linking AI use to responsible decision-making (Fadul et al., 2025; Martínez, 2026).

H3: Legal defensibility (LED) mediates the relationship between AI reliance (AIR) and physician accountability (ACC).

The results further demonstrate that legal defensibility (LED) significantly mediates the relationship between AI reliance (AIR) and physician accountability (ACC). The indirect effect (β = 0.250) is strong and statistically significant, with a high t-value of 8.048 and a p-value of 0.000, indicating a robust mediation effect. The confidence interval (0.201, 0.303) excludes zero, confirming the reliability of the effect. The VAF value of 54.47% suggests partial mediation with a relatively stronger mediating role compared to explainability, indicating that legal defensibility accounts for more than half of the total effect (β = 0.459). Although the direct effect (β = 0.209) remains significant, the stronger indirect pathway highlights the critical importance of legal justification in shaping accountability. These results suggest that increased reliance on AI systems necessitates stronger legal justification of clinical decisions, which in turn enhances physician accountability, supporting recent literature that emphasizes the growing role of legal and regulatory considerations in AI-assisted healthcare decision-making (Devetzis et al., 2025; Levy, 2026).

Table 4: Mediation Type and Effect

Hypothesis	Indirect Eff	t-val	P-V	CI (LL,	Direct Ef	Total Ef	VAF (	Mediation

	ect (β)	ue	alue	UL)	fect (β)	fect (β)	%)	Result
H2: AIR → EXP → ACC	0.120	5.078	0.000	(0.083, 0.160)	0.209	0.329	36.47%	Partial Mediation
H3: AIR → LED → ACC	0.250	8.048	0.000	(0.201, 0.303)	0.209	0.459	54.47%	Partial Mediation

Predive relevance, R square and effect size

The results presented in Table 5 demonstrate the predictive relevance, explanatory power, and effect sizes of the structural model across the key constructs, including accountability (ACC), explainability (EXP), and legal defensibility (LED). The R^2 value for ACC is 0.462, indicating that approximately 46.2% of the variance in physician accountability is explained by AI reliance (AIR), explainability, and legal defensibility, which can be considered moderate explanatory power in behavioral and healthcare research. Similarly, the R^2 values for EXP (0.317) and LED (0.421) suggest that AI reliance explains a substantial proportion of variance in these mediating constructs, confirming the model's adequacy in capturing key relationships. The Q^2 values for ACC (0.295), EXP (0.216), and LED (0.287) are all greater than zero, indicating strong predictive relevance of the model and its capability to accurately predict endogenous constructs. Furthermore, the effect size (f^2) values show that AIR has a small effect on ACC (0.039), a small effect on EXP (0.057), and a medium to large effect on LED (0.159), suggesting that AI reliance has a stronger influence on legal defensibility compared to other constructs. Overall, these findings confirm that the model demonstrates satisfactory predictive accuracy, explanatory strength, and meaningful effect sizes, supporting its robustness and suitability for analyzing complex relationships in AI-assisted clinical decision-making contexts (Hair et al., 2022; Sarstedt et al., 2025).

Table 5: Predive relevance R Square and effect size

	ACC	AIR	EXP	LED
ACC				
AIR	0.039		0.463	0.726
EXP	0.057			
LED	0.159			
R square	0.462	0.317	0.421	
Q Square	0.295	0.216	0.287	

Discussion

The findings of this study provide important insights into the evolving dynamics of physician accountability in AI-assisted clinical decision-making by empirically validating both direct and mediated relationships among AI reliance (AIR), explainability (EXP), legal defensibility (LED), and accountability (ACC). The significant direct effect of AI reliance on accountability supports the argument that physicians remain central decision-makers even in technologically augmented environments, reinforcing the principle that accountability is not transferred to AI systems but remains with human actors. This is consistent with recent medico-legal and governance literature, which emphasizes that clinicians retain ultimate responsibility for patient outcomes regardless of AI involvement, particularly under emerging regulatory frameworks (Khadke, 2026; Adebayo et al., 2026). Furthermore, the finding that increased AI reliance enhances accountability may reflect heightened awareness among physicians of the need to justify AI-assisted decisions, especially in high-risk clinical contexts. However, some opposing studies suggest that excessive reliance on AI may lead to automation bias and diffusion of responsibility, potentially weakening accountability when clinicians overly depend on algorithmic outputs without critical evaluation (Zhumanova et al., 2025). This apparent contradiction can be justified by the contextual nature of AI use, where accountability is strengthened when reliance is accompanied by oversight, but weakened when reliance becomes uncritical or passive.

The mediation results further deepen the understanding of how accountability is constructed through intermediate mechanisms. The significant mediating role of explainability (EXP) supports prior research highlighting that interpretability is essential for informed and responsible clinical decision-making, as it enables physicians to understand, evaluate, and justify AI-generated recommendations (Suryadevara et al., 2026; Kohan et al., 2026). This aligns with human-centered AI perspectives, which argue that explainability enhances trust calibration and preserves professional judgment. However, contrasting evidence suggests that explainability alone may not fully resolve accountability challenges, as simplified explanations can sometimes obscure uncertainty or create a false sense of understanding, thereby limiting their effectiveness in complex clinical scenarios (Sode, 2026). This justifies the partial mediation observed in the study, indicating that while explainability is important, it is insufficient on its own

to fully translate AI reliance into accountability. Similarly, the stronger mediating effect of legal defensibility (LED) underscores the growing importance of legal and regulatory considerations in AI-assisted healthcare, supporting recent studies that emphasize the need for decisions to be not only interpretable but also justifiable under legal scrutiny (Devetzis et al., 2025; Caputo, 2026). At the same time, opposing perspectives highlight ongoing ambiguity in legal frameworks and responsibility allocation, where accountability may be distributed across multiple stakeholders, including developers and institutions, rather than resting solely with physicians (Antonini, 2025). This tension explains why legal defensibility plays a dominant yet partial mediating role, as it strengthens accountability but does not fully eliminate the complexity of responsibility attribution in AI-enabled healthcare. Overall, the findings contribute to the literature by demonstrating that physician accountability is a multi-layered construct shaped by behavioral reliance, cognitive interpretability, and legal justification, thereby offering a more nuanced and realistic understanding of accountability in the age of clinical AI.

Theoretical Implications

The theoretical implications of this study lie in its advancement of a more integrated and multi-layered understanding of accountability in AI-assisted clinical decision-making by bridging gaps between technology adoption, explainability, and legal accountability theories. Unlike prior research that treats these constructs independently, this study contributes to theory by empirically validating a sequential and partially mediated framework linking AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC), thereby extending accountability theory and human–AI interaction literature into a medico-legal context. The findings support the argument that accountability is not a direct outcome of AI usage but is constructed through cognitive and legal mechanisms, reinforcing recent theoretical perspectives that emphasize layered accountability systems in AI governance (Khadke, 2026; Sode, 2026). Furthermore, by introducing legal defensibility as a mediating construct, the study addresses a critical gap in existing models that focus primarily on transparency and trust, thereby contributing to emerging scholarship that calls for integrating legal reasoning into AI accountability frameworks (Devetzis et al., 2025; Caputo, 2026). This enriches the conceptualization of explainable AI by positioning it as a necessary but insufficient condition for accountability, thus offering a more comprehensive theoretical model aligned with contemporary regulatory and ethical debates.

Practical Implications

From a practical perspective, the study provides valuable implications for healthcare practitioners, managers, policymakers, and AI developers by offering actionable insights into how accountability can be strengthened in AI-assisted environments. For clinicians, the findings highlight the importance of maintaining critical oversight and interpretive engagement with AI systems to ensure accountable decision-making, rather than relying passively on algorithmic outputs. For healthcare managers and administrators, the study underscores the need to design governance frameworks and clinical protocols that incorporate explainability and legal defensibility as core components of AI implementation, thereby reducing medico-legal risks and enhancing decision transparency (Adebayo et al., 2026; Kohan et al., 2026). Policymakers and regulators can benefit from the evidence by developing clearer guidelines that emphasize not only transparency but also legal justification and documentation of AI-assisted decisions, in line with evolving regulatory standards. Additionally, AI developers are encouraged to embed explainability and traceability features into system design to support both clinical usability and legal compliance. Overall, the study contributes to practice by demonstrating that effective AI integration in healthcare requires a balanced approach that aligns technological capability with professional accountability and legal robustness, ensuring safer and more trustworthy clinical decision-making.

Conclusion

In conclusion, this study provides a comprehensive synthesis of how physician accountability (ACC) is shaped within AI-assisted clinical decision-making by integrating behavioral, cognitive, and legal dimensions into a unified framework. The findings confirm that AI reliance (AIR) has a significant direct impact on accountability, reinforcing the notion that physicians remain ultimately responsible for clinical outcomes even in technologically augmented environments. At the same time, the study demonstrates that this relationship is not purely direct but is substantially influenced by intermediate mechanisms, particularly explainability (EXP) and legal defensibility (LED). Explainability enhances physicians' ability to interpret and justify AI-generated recommendations, thereby supporting accountable decision-making, while legal defensibility further strengthens this process by ensuring that such decisions can withstand regulatory and medico-legal scrutiny. The stronger mediating role of legal defensibility highlights the increasing

importance of legal justification in AI-integrated healthcare, reflecting the shift toward compliance-driven and accountability-centered clinical governance. Overall, the study advances the understanding that physician accountability in AI-assisted contexts is a multi-layered construct that emerges through the interaction of reliance, interpretability, and legal validation rather than from any single factor alone. The findings reconcile both supporting and opposing perspectives in the literature by showing that while AI reliance can enhance accountability when accompanied by oversight and justification, it may also pose risks if not supported by explainability and legal clarity. By empirically validating this integrated pathway, the study contributes to a more nuanced and realistic understanding of accountability in modern healthcare, emphasizing that effective AI adoption requires not only technological capability but also strong interpretive and legal frameworks. This synthesis underscores the need for balanced AI integration strategies that align clinical practice with ethical responsibility and legal robustness, ultimately ensuring safer, more transparent, and accountable healthcare delivery in the age of artificial intelligence.

Limitations of the Study: This study is subject to several limitations that should be acknowledged when interpreting the findings. First, the use of a cross-sectional research design limits the ability to establish causal inferences among AI reliance (AIR), explainability (EXP), legal defensibility (LED), and physician accountability (ACC), as the relationships are examined at a single point in time rather than across evolving clinical practices. Second, the reliance on self-reported data from healthcare professionals may introduce common method bias and subjective perceptions, particularly in assessing accountability and legal defensibility, which are context-sensitive and may vary across institutional settings. Third, the study focuses primarily on physicians within AI-enabled healthcare environments, which may limit generalizability to other stakeholders such as nurses, administrators, or AI developers who also play roles in accountability structures. Additionally, while the study introduces legal defensibility as a key construct, its measurement is still emerging and may not fully capture the complexity of legal reasoning and jurisdictional variations in AI governance frameworks (Mohammad, 2026). Furthermore, the model does not incorporate potential moderating variables such as organizational culture, regulatory environment, or level of AI system autonomy, which may influence the strength and direction of the observed relationships.

Future Research Directions

Future research should address these limitations by adopting longitudinal and experimental designs to better capture causal dynamics and changes in physician accountability over time as AI systems evolve. There is also a need for multi-stakeholder studies that include perspectives from policymakers, healthcare managers, patients, and AI developers to develop a more holistic understanding of accountability in AI-assisted healthcare ecosystems. Future studies could extend the model by incorporating moderating variables such as regulatory strictness, clinical risk level, and institutional governance mechanisms to examine boundary conditions of the relationships identified in this study. Moreover, further refinement and validation of the legal defensibility construct are necessary, particularly through interdisciplinary approaches that integrate legal scholarship and empirical healthcare research. Comparative studies across different countries and regulatory frameworks would also be valuable in understanding how legal and cultural contexts shape accountability in AI-assisted decision-making. Finally, future research should explore additional outcomes such as patient trust, clinical performance, and ethical compliance to expand the practical relevance of the model and contribute to the development of more robust, accountable, and human-centered AI systems in healthcare.

References:

Abbas, S. R., Seol, H., Abbas, Z., & Lee, S. W. (2025). Exploring the role of artificial intelligence in smart healthcare: a capability and function-oriented review. Healthcare,

Adhikari, K., Naik, N., Hameed, B. Z., Raghunath, S., & Somani, B. K. (2024). Exploring the ethical, legal, and social implications of ChatGPT in urology. *Current Urology Reports*, *25*(1), 1-8.

Ajmera, P., Gandikota, G., & Kharat, A. (2025). Artificial Intelligence in Hand and Wrist Imaging: Enhancing Diagnostics and Workflow. In *Imaging of the Hand and Wrist: Techniques and Applications* (pp. 571-585). Springer.

Alanazi, A. (2025). Assessing Clinicians' legal concerns and the need for a regulatory framework for AI in healthcare: a mixed-methods study. Healthcare,

Alrehaili, R. S., Almuzaini, S., Alrajhi, J., Dashti, A., Alsuroor, O., Alzahrani, F., Albishri, A., Almousa, M.,

Homdi, L., & Alshammakhy, R. (2025). Teledentistry in the Era of Digital Dentistry: Clinical Applications, Patient Experience, and Equity-Oriented Policy Implications. *Cureus*, *17*(12).

Arif, W. M. (2024). Radiologic technology students' perceptions on adoption of Artificial Intelligence technology in radiology. *International Journal of General Medicine*, 3129-3136.

Burmeister, J. R., Dimock, E., Haupert, M., & Zazay, I. (2025). Bridging the AI implementation gap in otolaryngology: A clinical commentary. *Digital Health*, *11*, 20552076251396981.

Camacho Clavijo, S. (2024). AI assessment tools for decision-making on telemedicine: liability in case of mistakes. *Discover Artificial Intelligence*, *4*(1), 24.

Chukwu, I. S., Onah, C. K. K., Nwankwo, E. P., & Ezomike, U. (2025). Ethical considerations and challenges in the use of artificial intelligence in pediatric surgical practice: a national survey of Nigerian pediatric surgeons. *World Journal of Pediatric Surgery*, *8*(5), e001089.

Cohen, I. G., & Topol, E. J. (2022). Algorithmic Accountability in AI-Driven Diagnostic Platforms: Bridging Clinical Safety Standards and Legal Liability Models.

Cooney-Waterhouse, T., Ou, W., Mukherji, S., Frytak, J., Saha, P., & Waterhouse, D. (2025). Bridging the Trust Gap in Artificial Intelligence for Health care: Lessons from Clinical Oncology. *AI in Precision Oncology*, *2*(3), 85-91.

Coppola, F., Faggioni, L., Gabelloni, M., De Vietro, F., Mendola, V., Cattabriga, A., Cocozza, M. A., Vara, G., Piccinino, A., & Lo Monaco, S. (2021). Human, all too human? An all-around appraisal of the "artificial intelligence revolution" in medical imaging. *Frontiers in Psychology*, *12*, 710982.

Dai, T., & Singh, S. (2025). Artificial intelligence on call: the physician's decision of whether to use AI in clinical practice. *Journal of Marketing Research*, *62*(5), 854-875.

Dalky, A., Osaid Malkawi, R. M., Alrawashdeh, A., ALBashtawy, M., & Hani, S. B. (2025). Perceptions, barriers, and risk of artificial intelligence Among healthcare professionals: A cross-sectional study. *Digital Health*, *11*, 20552076251360924.

Damiani, G., Altamura, G., Zedda, M., Nurchis, M. C., Aulino, G., Heidar Alizadeh, A., Cazzato, F., Della Morte, G., Caputo, M., & Grassi, S. (2023). Potentiality of algorithms and artificial intelligence adoption to improve medication management in primary care: a systematic review. *BMJ open*, *13*(3), e065301.

De Mukhopadhyay, K., Das, T., & Basu, A. (2025). Explainable AI in Oral Carcinoma Detection: Legal Challenges and Policy Pathways in India. 2025 International Conference on Artificial Intelligence for Computing, Astronomy and Renewable Energy (AICARE),

De Simone, B., Deeken, G., & Catena, F. (2025). Balancing ethics and innovation: can artificial intelligence safely transform emergency surgery? A narrative perspective. *Journal of clinical medicine*, *14*(9), 3111.

Dehouwer, R., Vernieuwe, L., & Versyck, B. (2025). From Probe to Practice: Implementing Point-of-Care Ultrasound in Anesthesiology. *Best Practice & Research Clinical Anaesthesiology*.

Di Mauro, L., Capasso, E., Tettamanti, C., Casella, C., Francaviglia, M., Volonnino, G., Rinaldi, R., Esposito, M., & Chisari, M. (2025). The role of artificial intelligence in analyzing clinical malpractice disputes through medical record management. *Journal of Forensic and Legal Medicine*, 102941.

Di Palma, G., Scendoni, R., De Benedictis, A., Tambone, V., & De Micco, F. (2025). Leveraging artificial intelligence for collaborative care planning: Innovations and impacts in shared decision-making–A systematic review. *Open Medicine*, *20*(1), 20251232.

Dost, B., Turan, E. İ., Aydın, M. E., Ahıskalıoğlu, A., Narayanan, M., Yılmaz, R., & De Cassai, A. (2025). Artificial intelligence in anaesthesiology: current applications, challenges, and future directions. *Turkish Journal of Anaesthesiology and Reanimation*, *53*(6), 282.

Duan, L., Yao, Z., Li, X., Wu, Y., & Sheng, D. (2025). Comparing large language models and human doctors in symptom-driven online medical consultations: A case study on trigeminal neuralgia. *Digital Health*, *11*, 20552076251388140.

Granata, V., Fusco, R., Setola, S. V., Santorsola, M., Ottaiano, A., Cerrone, M., Pizza, N., Ferrara, G., Zerunian, M., & Caruso, D. (2025). An Update of AI and Radiomics in Precision Oncology: Insights from Liver Tumors as Case Models. *Technology in Cancer Research & Treatment*, *24*, 15330338251387928.

Hadjiathanasiou, A., Goelz, L., Muhn, F., Heinz, R., Kreißl, L., Sparenberg, P., Lemcke, J., Schmehl, I., Mutze, S., & Schuss, P. (2025). Artificial intelligence in neurovascular decision-making: a comparative analysis of ChatGPT-4 and multidisciplinary expert recommendations for unruptured intracranial aneurysms. *Neurosurgical Review*, *48*(1), 261.

He, F., & Chen, D. (2025). Expert review: current applications and future directions of artificial intelligence in obstetrics. *Gynecology and Obstetrics Clinical Medicine*, *5*(4).

Hosseini, F. (2025). Artificial intelligence in healthcare: applications, challenges, and future directions. A narrative review informed by international, multidisciplinary expertise.

Hui, L., Raff, D., Hu, R., Yau, O., & Singla, R. (2025). Artificial intelligence in family medicine: Opportunities, impacts, and challenges: Integrating artificial intelligence into family medicine will have impacts across the continuum of patient care. *British Columbia*

Medical Journal, *67*(2).
Jairoun, A. A., Al-Hemyari, S. S., Shahwan, M., Al-Qirim, T., & Shahwan, M. (2024). Benefit–risk assessment of chatgpt applications in the field of diabetes and metabolic illnesses: a qualitative study. *Clinical Medicine Insights: Endocrinology and Diabetes*, *17*, 11795514241235514.
Katta, M. R., Kalluru, P. K. R., Bavishi, D. A., Hameed, M., & Valisekka, S. S. (2023). Artificial intelligence in pancreatic cancer: diagnosis, limitations, and the future prospects—a narrative review. *Journal of Cancer Research and Clinical Oncology*, *149*(9), 6743-6751.
Lawrence, K., Kuram, V. S., Levine, D. L., Sharif, S., Polet, C., Malhotra, K., & Owens, K. (2025). Informed consent for ambient documentation using generative AI in ambulatory care. *JAMA network open*, *8*(7), e2522400.
Mandelia, A., Rengan, V. S., Mehta, A. R., Vankam, A. V., Babu, R., Biswas, S. K., & Agrawal, V. (2025). Artificial intelligence use in daily and professional life among pediatric surgeons in India: A roadmap for adoption based on online survey results. *Journal of Indian Association of Pediatric Surgeons*, *30*(3), 361-368.
Margalit, Y., & Shomron, N. (2025). From Responsible Genomics to Responsible AI in Healthcare.
Mensah, G. B. (2024). Ensuring AI Algorithm Fairness in Healthcare Decision-Making. In: Research-Gate.
Mi, Y., O'Reilly, B. A., & Tabirca, S. (2024). Medical software as a virtual service: A multifaceted approach to telemedicine through software-as-a-service and user-centric features. *Digital Health*, *10*, 20552076241295441.
Musa, Y. (2025). The Expanding Role of Artificial Intelligence in Clinical Medical Radiology Practice. *Annals of Medical and Health Research: An International Journal. DOI: https://doi. or g/10.51470/ARMHR*, *1*.
Nadeem, M., Alzoubi, Y. I., Elbasi, E., & Topcu, A. E. (2025). Robot-Assisted Surgery: A Comprehensive Review of Literature, Challenges, Ethical Considerations, and Current Trends. *IEEE Access*, *13*, 215647-215680.
Naidu, G., & Krishnan, V. (2025). Artificial Intelligence-Powered Legal Document Processing for Medical Negligence Cases: A Critical Review. *Int. J. Intell. Sci*, *15*(1), 10-55.
Nair, S. C., Boma, N., Ravindran, P., Shah, C., Joseph, L., Satish, K. P., & Antony, E. (2025). The Bridge "Artificial Intelligence" Over the "Healthcare Quality" Lake. *International Journal for Healthcare Quality, Patient Centeredness & Safety*, *6*(1), 2-3.
Nzeako, T. R., Elendu, C., Echefu, G., Olanisa, O., Kiladejo, A., & Bob-Manuel, E. D. (2025). Artificial intelligence in interventional cardiology: a review of its role in diagnosis, decision-making, and procedural precision. *Annals of Medicine and Surgery*, *87*(9), 5720-5734.
O'sullivan, S., Janssen, M., Holzinger, A., Nevejans, N., Eminaga, O., Meyer, C., & Miernik, A. (2022). Explainable artificial intelligence (XAI): closing the gap between image analysis and navigation in complex invasive diagnostic procedures. *World Journal of Urology*, *40*(5), 1125-1134.
Okoye, O., Nwachukwu, N., Uche, N., Udeh, N., & Umeh, R. (2024). AFRICAN JOURNAL OF BIOETHICS.
Olawade, D. B., Ojo, I. O., Oisakede, E. O., Joel-Medewase, V. I., & Wada, O. Z. (2025). Artificial Intelligence in Nigerian Oncology Practice: A Qualitative Exploration of Oncologists' Perspectives. *Journal of Cancer Policy*, 100626.
Parizad, R., Hatwal, J., Javanshir, E., Batta, A., & Mohan, B. (2025). Artificial Intelligence in Cardiopulmonary Resuscitation: Revolutionizing Resuscitation Through Precision and Prediction–A Narrative Review. *Vascular Health and Risk Management*, 847-857.
Polizzi, A., Serra, S., Leonardi, R., & Isola, G. (2025). Clinical applications of artificial intelligence in teleorthodontics: a scoping review. *Medicina*, *61*(7), 1141.
Quon, S., Low, S., Zhou, S., & Zheng, K. (2025). Medical Education on Provider-Patient Power Dynamics: A Review of the Literature on Training for Responding to Patient Reports of Physician Misconduct. *Journal of Medical Education and Family Medicine*, *2*(3), 108-117.
Ramessur, R., Raja, L., Kilduff, C. L., Kang, S., Li, J.-P. O., Thomas, P. B., & Sim, D. A. (2021). Impact and challenges of integrating artificial intelligence and telemedicine into clinical ophthalmology. *The Asia-Pacific Journal of Ophthalmology*, *10*(3), 317-327.
Raza, A. (2024). Navigating the intersection of artificial intelligence and law in healthcare: complications and corrections. *Policy Review Journal*, *2*, 2364.
Sachdeva, C., & Jain, M. P. (2025). Ethical Considerations of Using Artificial Intelligence and Associated Legal Frameworks in Tertiary Healthcare Organisations. *INTERNATIONAL JOURNAL OF ENGINEERING DEVELOPMENT AND RESEARCH*, *13*(1), 97-105-197-105.
Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare,
Sallam, M. (2025). AI at the crossroads of cost and care: A SWOT analysis of microsoft AI diagnostic orchestrator (MAI-DxO) in clinical diagnosis. *Journal of Clinical Medicine and Research*.
Sangers, T. E., Kittler, H., Blum, A., Braun, R. P., Barata, C., Cartocci, A., Combalia, M., Esdaile, B., Guitera, P., & Haenssle, H. A. (2024). Position statement of the EADV artificial intelligence (AI) task force on AI-assisted smartphone apps and web-based services for

skin disease. *Journal of the European Academy of Dermatology and Venereology*, *38*(1), 22-30.

Sengul, T., Sariköse, S., & Gul, A. (2025). Ethical decision-making and artificial intelligence in nursing education: An integrative review. *Nursing ethics*, *32*(8), 2490-2515.

Serban, I. V., & Dan, A. N. (2025). The New Frontiers of Socialised Medicine: AI, Bio-Legal and Clinical-Legal Liability Between EU Law and the Italian Legal System. *BioLaw Journal-Rivista di BioDiritto*(3S), 55-67.

Sevgi, K., Pliskow, S., & Aran, S. N. (2025). Pelvic Fractures in Pregnancy: Multidisciplinary Management and Outcomes. *Cureus*, *17*(5).

Shah, R., Bozic, K. J., & Jayakumar, P. (2025). Artificial intelligence in value-based health care. *HSS Journal®*, *21*(3), 307-313.

Singh, G., & Bansal, M. (2025). Role of Artificial Intelligence in Anesthesia. *European Journal of Cardiovascular Medicine*, *15*(9).

Taghipour, N., Mostafavinia, A., Hosseini, S., Javaherinasab, M., Mehran, H., Alighadr, A., Javankiani, S., & Dehbozorgi, S. (2025). Surgeon Attitudes Toward AI-Enhanced Robotic Surgery: A Survey on Autonomy, Liability, and Skill Erosion. *InfoScience Trends*, *2*(6), 29-41.

Tukur, H. N., Uwishema, O., Akbay, H., Sheikhah, D., & Correia, I. F. S. (2025). AI-assisted ophthalmic imaging for early detection of neurodegenerative diseases. *International Journal of Emergency Medicine*, *18*(1), 90.

Venturi, F., Veronesi, G., Gualandi, A., Magnaterra, E., Scotti, B., Sotiri, I., Baraldi, C., Alessandrini, A. M., Veneziano, L., & Vaccari, S. (2025). From Slide to Insight: The Emerging Alliance of Digital Pathology and AI in Melanoma Diagnostics. *Cancers*, *17*(22), 3696.

Veparala, A. S., Lall, D., Srinivas, P. N., Samantaray, K., & Marchal, B. (2025). Health system drivers of caesarean deliveries in south Asia: a scoping review. *The Lancet Regional Health-Southeast Asia*, *40*.

Wu, R., Ge, E., Huang, B., Bi, Z., Wang, T., Hao, J., & Wu, Z. (2025). From Correlation to Causation: The Foundations of Trustworthy Clinical AI. *Available at SSRN 5739384*.

Zhang, J., Li, D., Wang, X., Wu, Z., & Lyu, Q. (2025). Application and challenges of DeepSeek in primary care in China. *Frontiers in public health*, *13*, 1721308.