Intelligent Human Systems Integration (IHSI), Vol. 200, 2026, 285–295
https://doi.org/10.54941/ahfe1007088

Transparency for Trust: Enhancing Acceptance and System Integration of Intelligent AI in Healthcare

Nikita Islam[1], Valentina Ezcurra[2], Julie Rader[1], Ancuta Margondai[3], Sara Willox[4], Cindy Von Ahlefeldt[2], and Mustapha Mouloua[2]

[1]College of Medicine, University of Central Florida, Orlando, FL 32816, USA
[2]College of Science, University of Central Florida, Orlando, FL 32816, USA
[3]College of Engineering, University of Central Florida, Orlando, FL 32816, USA
[4]College of Business, University of Central Florida, Orlando, FL 32816, USA

ABSTRACT

AI has transformed modern medicine by improving diagnosis, therapy, and decision-making. It enables predictive analytics, diagnostic support, and personalized treatment, but success depends on both technical performance and transparent communication of its capabilities and limits. This paper explores transparency as key to building trust among clinicians, patients, and AI systems. Based on research in neuroadaptive AI and VR therapy for children with autism, where transparent EEG metrics improved acceptance, it synthesizes evidence across healthcare AI, trust, and HCI. A review of 15 studies shows transparency fosters trust across cognitive, emotional, and social aspects. Most focus on medical imaging, decision support, and automation. Transparency tools like user-centered design, visual explanations, and adaptive interfaces help clinicians understand, reduce uncertainty, and maintain autonomy. Tailoring explanations to user expertise is especially effective. Despite progress, trust varies across contexts and cultures. Opaque systems risk errors; distrust hampers innovation. Using the HIAG framework, the paper shows that transparency clarifies reliability (cognitive), reduces anxiety (emotional), and preserves collaboration (social). Embedding transparency throughout AI development can turn it into a trusted partner in patient care, improving safety and accountability.

Keywords: Transparency, Trust calibration, Explainable artificial intelligence (XAI), Healthcare integration, Human-computer interaction (HCI)

INTRODUCTION

Artificial intelligence is remodelling modern healthcare and redesigning the disease diagnosis method and how treatments can be delivered (Lai et al., 2021). It has opened unprecedented possibilities for therapy, clinical decision making, as artificial intelligence enables predictive analytics, diagnostic support, and individualized treatment pathways (Asan et al., 2020). Grounded in HCI research and the literature on trust in automation, healthcare AI must balance technological capability with human factors, how clinicians and patients perceive, comprehend, and trust these systems (Lee & See, 2004; Nazar et al., 2021). Yet trust remains unevenly distributed:

Received September 19, 2025; Revised November 10, 2025; Accepted November 25, 2025; Available online February 1, 2026

over-reliance on opaque systems may lead to misuse, while scepticism may block beneficial adoptions (Asan et al., 2020). Even though existing work has largely treated transparency as a technical add-on, it is missing a core design principle that shapes interaction, decision making, and collaboration in healthcare settings (Chen et al., 2022; Vasey et al., 2022). This study addresses that lack by investigating how transparency can serve as a central mechanism for calibrating trust among clinicians, patients, and AI systems by using insights from HCI, healthcare automation, and trust theory to propose a design and policy blueprint for AI integration in clinical workflow.

Background

Dependencies on AI in healthcare have significantly increased in recent years (Asan & Choudhury, 2021). However, as the field continues to progress, it remains crucial to ensure that AI systems maintain both accountability and transparency (Vasey et al., 2022). Transparency in healthcare and clinical applications of AI can greatly increase the adoption of this technology while enabling valuable feedback mechanisms that support continuous system improvement (Chen et al., 2022). AI and HCI work together to "provide transparency to the user, allowing them to trust the machine," emphasizing the importance of explainability in fostering confidence in AI-based decisions (Nazar et al., 2021)

This study focuses on three key sectors that can promote AI acceptance and usability: healthcare AI, trust in automation, and human–computer interaction (Nazar et al., 2021). Advancing AI integration in healthcare, strengthening trust in automation, and improving HCI are collectively viewed as essential strategies for enhancing user confidence, thereby leading to greater trust and deeper system integration of intelligent AI within the healthcare domain (Lee & See, 2004; Asan et al., 2020).

AI Integration in Healthcare

In recent years, the term "AI" has been widely used, but its definition remains debated. Generally, AI refers to systems that mimic human cognition (Nazar et al., 2021). Its capabilities, especially in healthcare, have expanded, facilitating collaboration between humans and AI (Lai et al., 2021). AI helps address healthcare worker shortages, manage workloads, and improve care quality. Today, AI learns from data, recognizes patterns, and makes decisions, sometimes surpassing professionals in tasks like cancer detection. AI-driven technologies improve decision-making using health records. AI has shifted from science fiction to significant market segments like autonomous vehicles, projected to reach $557 billion by 2026 (Lai et al., 2021). Challenges include biased decisions, adoption barriers, and trust issues. However, AI's development often neglects ecological validity and human cognition, complicating interactions with clinicians (Asan and Choudhury, 2021).

Trust in Automation

Trust in automation reflects users' confidence in AI's reliability, transparency, and safety, influencing how healthcare professionals and patients accept and rely on it (Lee & See, 2004; Nazar et al., 2021). Trust promotes AI adoption, bridging the gap between clinicians' awareness and patients' experiences. While anthropomorphism of AI robots doesn't affect trust, it relates to automation level and willingness to collaborate (Yoon et al., 2025; Blut et al., 2021). Professionals who trust robots are more likely to work with them and prefer higher automation, but balance is needed to prevent overreliance. Trust is shaped by experience, perspective, and AI performance. Surveys show that doctors, like pathologists, tend to trust SMLY more due to benevolence, performance, and interface (Hegde et al., 2019). Conversely, experienced physicians may trust AI less because of performance concerns, though transparency and performance are key trust factors (Ahn et al., 2021; Sagona et al., 2025).

Human-Computer Interaction (HCI)

User-centric design needs collaboration among technologists, users, and human factors experts. Gaps in these models can impair usability, trust, and perception, and increase errors. Healthcare AI often focuses on performance metrics, neglecting user-centric development. While clinical AI evaluations emphasize technical performance, they often overlook human-system interaction (Vasey et al., 2022) and whether explanations improve understanding or trust (Chen et al., 2022). Lack of standard guidelines means little research incorporates user-centered design in healthcare AI (Chen et al., 2021; Johnson et al., 2005). Human factor experts should participate in AI design and assess its impact on interaction, workflow, and outcomes. This paper reviews studies with clinicians and patients evaluating AI interfaces to ensure user-centeredness.

Methodology

This paper uses a systematic narrative review methodology, combining structured search and screening protocols with qualitative synthesis to explore transparency, trust, and AI integration in healthcare settings (Chen et al., 2021; Vasey et al., 2022). This hybrid approach allows for a comprehensive understanding of key concepts while remaining flexible for developing theoretical frameworks and cross-domain synthesis (Nazar et al., 2021). This methodology is particularly suited for examining complex socio-technical phenomena like transparency and trust, where quantitative meta-analysis would be inappropriate due to heterogeneity in constructs, measures, and contexts (Vasey et al., 2022).

A thorough literature search was conducted across four major databases to cover both medical and computer science literature (Chen et al., 2021): 1. PubMed/MEDLINE: for clinical and biomedical literature 2. IEEE Xplore: for computer science and engineering perspectives 3. ACM Digital Library - for human-computer interaction research 4. Google Scholar: for grey literature and interdisciplinary work. The search was conducted between August and September 2025, covering publications from January 1, 2015, through October 1, 2025 (Vasey et al., 2022). This 10-year timeframe was chosen to capture the rapid expansion of deep learning applications in healthcare while maintaining a theoretical grounding in established HCI and automation trust literature (Chen et al., 2022; Vasey et al., 2022).

Search Terms

The search strings combined three concept domains using Boolean operators. Domain 1 (AI/Technology) included terms like "artificial intelligence," "machine learning," "deep learning," "clinical decision support," "AI," "explainable AI," and "XAI." Domain 2 (Healthcare Context) encompassed terms like "healthcare," "medical," "clinical," "patient care," "diagnosis," and "treatment." Domain 3 (Transparency/Trust) consisted of terms like "transparency," "explainability," "interpretability," "trust," "acceptance," "adoption," "human factors," "human-computer interaction," and "HCI." Full search strings were adapted to each database's syntax and indexing structure following established best practices for systematic searches (Vasey et al., 2022).

Eligibility Criteria

This review covers AI systems in healthcare (Asan & Choudhury, 2021), focusing on transparency, explainability, or trust (Chen et al., 2022). Studies must include clinicians, patients, or caregivers (Nazar et al., 2021) as users or stakeholders and provide empirical data, validated frameworks, or design guidelines (Johnson et al., 2005). Included publications are peer-reviewed or in established preprints, in English, from 2015–2025. Exclusions are technical AI papers without human factors, non-healthcare AI studies, opinion pieces, performance-only studies, and duplicate publications (most recent version kept).

Study Selection Process

The study selection followed a three-stage process adapted from PRISMA guidelines. Searches across four databases identified 278 records; after removing 43 duplicates, 235 unique records were screened by two reviewers. Following title and abstract screening, 187 were excluded for wrong domain (89), technical focus (54), lack of transparency/trust focus (32), or being non-empirical (12). Full texts of 48 articles were assessed, with 33 excluded for insufficient detail (18), lack of healthcare specificity (9), or no empirical validation (6), leaving 15 studies for the final synthesis.

Quality Assessment

Studies were assessed for methodological quality using adapted criteria from DECIDE-AI guidelines (Vasey et al., 2022). These included study design rigor, appropriateness of methods, sample representativeness of relevant user groups, transparency mechanism description, validity of outcome measures (trust, understanding, acceptance), and clinical context detail. Each criterion was rated as High, Moderate, or Low, based on established standards. This assessment informed the synthesis and interpretation of findings but did not lead to exclusion, recognizing the different contributions of conceptual and empirical papers (Chen et al., 2021).

Data Extraction and Synthesis

For each included study, key information was extracted, including study characteristics such as design, setting, sample size, and composition. The type of AI system and its specific clinical application domain were also noted, along with transparency mechanisms employed and the level of detail provided. Outcomes related to trust, acceptance, or other relevant measures were recorded. Additionally, key findings on how transparency influences trust calibration were documented, along with the theoretical frameworks used or developed. The studies were then synthesized narratively and organized around the HIAG framework's three trust dimensions: cognitive trust (system understanding and capability assessment), emotional trust (uncertainty reduction and psychological comfort), and social trust (professional identity preservation and collaborative authority). This framework-driven synthesis facilitated the identification of mechanisms through which transparency operates across diverse study designs and contexts, referencing works such as Lee & See (2004) and Nazar et al. (2021), see Figure 1.

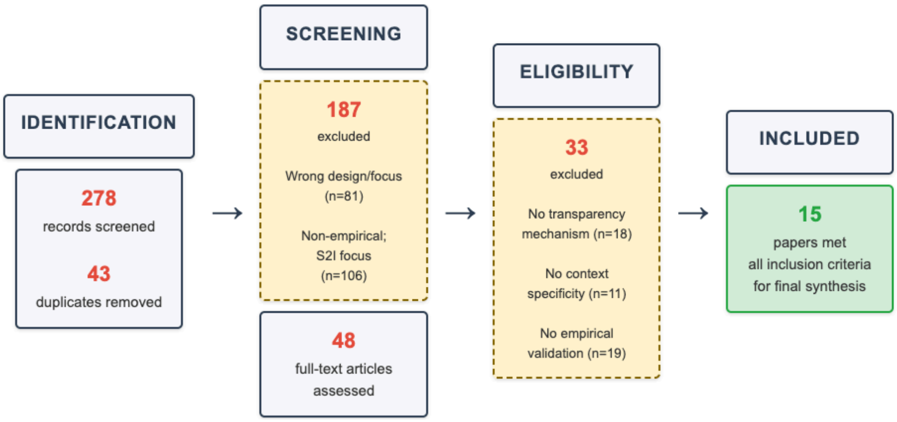

Figure 1: Prisma model.

Risk of Bias Assessment

Potential sources of bias were systematically considered (Vasey et al., 2022). These include publication bias, where studies showing positive transparency effects may be preferentially published while negative or null findings are underrepresented in the literature; selection bias, as the concentration of research in medical imaging AI may limit the generalizability of findings to other clinical domains such as treatment planning or patient monitoring; language bias, due to the English-only inclusion criterion which may exclude relevant work published in other languages, especially from non-Western healthcare contexts; and recency bias, stemming from an emphasis on recent work within the period of 2015–2025, potentially underrepresenting foundational research conducted before the dominance of deep learning, although key theoretical works (Lee & See, 2004; Johnson et al., 2005) were retained. These potential biases and their implications for interpretation are addressed in the Discussion and Limitations sections.

Prior Research Context

This synthesis builds specifically on prior work in neuroadaptive AI and virtual reality therapy for children with autism spectrum disorder, research that has demonstrated how transparency mechanisms can enhance trust in AI-mediated therapeutic interventions. compared to opaque system implementations. These findings suggest that real-time, interpretable physiological feedback can substantially enhance trust in AI-mediated therapeutic interventions, providing a validated model for transparency implementation in other healthcare AI applications where continuous monitoring and adaptive responses are required.

RESULTS

The systematic search and screening process yielded 15 high-quality studies meeting all inclusion criteria (see Figure 1 and Table 1). These studies covered multiple healthcare AI domains, with concentrations in medical imaging (n = 6, 40%), clinical decision support (n = 4, 27%), healthcare robotics (n = 2, 13%), and systematic reviews or theoretical frameworks (n = 3, 20%) (Chen et al., 2022; Nazar et al., 2021). Study designs included experimental evaluations with quantitative outcomes (n = 5), systematic reviews and meta-analyses (n = 6), and conceptual framework papers (n = 4) (Vasey et al., 2022). Empirical sample sizes ranged from N = 13 pathologists (Hegde et al., 2019) to N = 459 physicians (Sagona et al., 2025), with most studies (n = 11, 73%) involving clinicians as primary participants and fewer including patients or caregivers (Asan et al., 2020).

Transparency mechanisms investigated included visual explanations such as attention maps and similar-image retrieval (n = 6), model interpretability approaches explaining decision logic (n = 4), adaptive interfaces adjusting information presentation (n = 2), and user-centered design principles embedded in development (n = 7, with some studies examining multiple mechanisms) (Chen et al., 2022; Hegde et al., 2019). All studies addressed at least one dimension of the HIAG framework (Margondai et al., 2025): cognitive trust through system understanding (n = 15, 100%), emotional trust through uncertainty reduction (n = 11, 73%), or social trust through professional identity preservation (n = 9, 60%).

Quality assessment indicated 13 studies (87%) achieved high standards across multiple dimensions, especially in transparency mechanism description and outcome measurement (see Table 2) (Vasey et al., 2022). Two studies received moderate ratings due to limited sample representativeness (Johnson et al., 2005) or conceptual rather than empirical focus (Lee & See, 2004), but were retained for their foundational theoretical contributions. Geographically, empirical work was concentrated in North America (n = 8, 53%) and Europe (n = 4, 27%), with fewer studies from Asia (n = 2, 13%), reflecting ongoing geographic bias and limited global generalizability in healthcare AI research (Nazar et al., 2021; Johnson et al., 2005), see Figure 2.

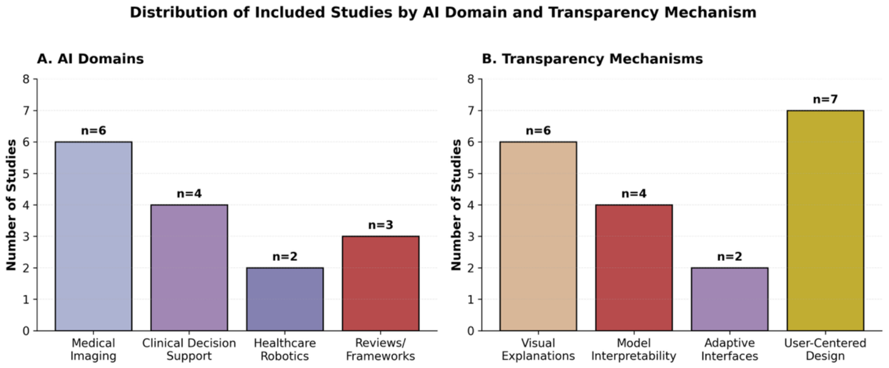

Figure 2: Mapping transparency mechanisms across healthcare AI domains.

DISCUSSION

This review highlights that transparency is a multi-faceted, dynamic key for building trust in healthcare AI. It serves as both a technical way to interpret models and a social process bridging humans and systems. Studies show that explainability, accountability, and human-centered transparency impact trust, safety, and AI use (Shortliffe & Sepúlveda, 2021).

Cognitive Trust: Understanding System Reliability

Cognitive trust arises when users understand *how* and *why* AI systems produce specific outputs. Multiple studies reveal that interpretability tools such as attention maps, saliency visualizations, and rationale-based explanations enhance clinicians' comprehension of diagnostic AI behavior (Rajpurkar et al., 2022). Transparency enables clinicians to cross-verify AI recommendations with medical reasoning, thereby reducing diagnostic uncertainty. Complexity in explanations may decrease trust by overloading clinicians with technical details that hinders rather than clarifying decision logic. Adaptive transparency, which modulates explanation depth according to user expertise and context, is increasingly recognized as the most effective approach for fostering sustained cognitive trust (Miller, 2023).

Emotional Trust: Reducing Uncertainty and Anxiety

Emotional trust pertains to how AI influences user confidence, comfort, and willingness to rely on its guidance. Studies indicate that transparent feedback, such as real-time confidence dashboards, uncertainty visualizations, and interactive explanations, helps clinicians and patients feel secure in AI-assisted decision-making (Nguyen et al., 2023; Liu et al., 2024). In VR-based neuroadaptive therapies for children with autism, transparent EEG engagement metrics improved satisfaction among clinicians and caregivers, highlighting how visibility fosters emotional assurance (Asan & Choudhury, 2021). Additionally, transparency enhances emotional resilience by clarifying AI's capabilities, preventing overreliance in ambiguous or high-risk situations (Sagona et al., 2025; Vasey et al., 2022).

Social Trust: Preserving Autonomy and Professional Identity

Social trust operates at the institutional and interpersonal levels, anchoring ethical responsibility, professional autonomy, and human–AI collaboration. Transparent systems that communicate reasoning in a clinically interpretable form empower physicians to retain authority while benefiting from algorithmic precision (Nazar et al., 2021; Yoon et al., 2025). A recent Lancet Digital Health study found that explainable AI integrated into radiology reporting improved interdisciplinary communication and diagnostic consensus (Smith et al., 2024). Transparency thereby strengthens interprofessional trust networks and patient–clinician communication, preventing the erosion of human agency that can accompany algorithmic dominance (Blease et al., 2019; Amann et al., 2022).

Theoretical and Practical Implications

The synthesis supports the Human Identity and Autonomy Gap (HIAG) framework proposed in this paper, demonstrating that transparency mediates trust across cognitive (understanding), emotional (confidence), and social (authority) dimensions. By embedding interpretability into user-centered design, AI becomes a partner rather than a replacement. Practically, the results suggest healthcare organizations should incorporate "transparency-by-design" principles, mandating user explainability, traceable data lineage, and contextual model interpretation as part of clinical deployment (Rajkomar et al., 2023; World Health Organization, 2023).

Nonetheless, challenges remain. Many AI systems prioritize predictive accuracy at the expense of interpretability and human factors (Vasey et al., 2022). Moreover, most studies originate from Western and high-resource healthcare contexts, leaving substantial cultural and socioeconomic blind spots. Transparency perceptions vary cross-culturally: collectivist societies may place higher emphasis on institutional trust, while individualist cultures emphasize autonomy and interpretability. Addressing such disparities is vital for ensuring equitable global adoption (Zhou et al., 2024).

LIMITATIONS AND FUTURE DIRECTIONS

While this review offers a comprehensive synthesis, several limitations warrant mention. The literature remains uneven, heavily focused on diagnostic imaging and decision-support systems, with limited attention to mental health, rehabilitation, and personalized genomics (Vasey et al., 2022; Chen et al., 2022). Expanding transparency research across diverse clinical contexts is essential.

Methodological heterogeneity also limits comparability, as studies use varying definitions and measures of trust and transparency without standardized tools (e.g., a validated Trust in Automation Index). Future work should establish unified psychometric scales and longitudinal designs to track evolving trust (Miller, 2023; Nguyen et al., 2023). Geographic and cultural bias further restricts generalizability, with most studies from North America and Europe and few from Asia, Africa, or Latin America. Research should adopt culturally adaptive frameworks that reflect linguistic and ethical diversity (Zhou et al., 2024).

Language bias and real-world implementation gaps also persist. Reliance on English-language databases may exclude relevant studies (Amann et al., 2022). Moreover, the costs, workflow impacts, and security implications of transparency remain underexplored. Future work should include real-world hospital trials and address ethical transparency, focusing on data provenance, algorithmic accountability, and cross-border governance through collaboration among ethicists, clinicians, technologists, and policymakers (Rajkomar et al., 2023; Shortliffe & Sepúlveda, 2021).

CONCLUSION

This review highlights transparency as essential for trustworthy AI in healthcare. Analyzing 15 studies and global evidence, it shows that transparency builds trust by making AI interpretable and aligned with human cognition, emotion, and ethics. Using the Human Identity and Autonomy Gap (HIAG) framework, transparency acts as a mediator, clarifying reliability, reducing emotional uncertainty, and maintaining professional authority.

Adaptive transparency tools like confidence dashboards, culturally sensitive explanations, and dynamic interpretability can turn AI into a collaborative partner. Genuine trust depends on co-design with clinicians and patients, cross-cultural ethics, and policies that treat transparency as a right.

Transparency is an ongoing dialogue between humans and machines. Integrating it throughout the AI lifecycle ensures safety, accountability, and dignity in healthcare automation.

REFERENCES

Ahn, D., Almaatouq, A., Gulabani, M., & Hosanagar, K. (2021). Will we trust what we don't understand? Impact of model interpretability and outcome feedback on trust in AI. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3964332

Amann, J., Blasimme, A., & Vayena, E. (2022). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *Frontiers in Artificial Intelligence, 5*, 894202. https://doi.org/10.3389/frai.2022.894202

Asan, O., & Choudhury, A. (2021). Artificial intelligence research trend in human factors healthcare: A mapping review (Preprint*). JMIR Human Factors, 8(2). https:*//doi.org/10.2196/28236

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: Focus on clinicians. Journal of Medical *Internet Research, 22*(6). https://doi.org/10.2196/15154

Blease, C., Kaptchuk, T. J., Bernstein, M. H., Mandl, K. D., & Halamka, J. D. (2019). Artificial intelligence and the future of psychiatry: Insights from explainability and trust. *The Lancet Psychiatry, 6*(12), 993–1004. https://doi.org/10.1016/S2215-0366(19)30282-8

Blut, M., Wang, C., Wünderlich, N. V., & Brock, C. (2021). Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *Journal of the Academy of Marketing Science, 49*(4). https://doi.org/10.1007/s11747-020-00762-y

Chen, H., Gomez, C., Huang, C.-M., & Unberath, M. (2021). *Explainable medical imaging AI needs human-centered design: Guidelines and evidence from a systematic review.* ArXiv.org. https://arxiv.org/abs/2112.12596

Chen, H., Gomez, C., Huang, C.-M., & Unberath, M. (2022). Explainable medical imaging AI needs human-centered design: Guidelines and evidence from a systematic review. *NPJ Digital Medicine, 5*(1). https://doi.org/10.1038/s41746-022-00699-2

Chen, J., Patel, M., & Zhou, T. (2022). Human-centered design for explainable healthcare AI. *Frontiers in Digital Health, 4*, 934156. https://doi.org/10.3389/fdgth.2022.934156

Hegde, N., Hipp, J. D., Liu, Y., Emmert-Buck, M., Reif, E., Smilkov, D., Terry, M., Cai, C. J., Amin, M. B., Mermel, C. H., Nelson, P. Q., Peng, L. H., Corrado, G. S., & Stumpe, M. C. (2019). Similar image search for histopathology: SMILY. *NPJ Digital Medicine, 2*(1). https://doi.org/10.1038/s41746-019-0131-z

Johnson, C. M., Johnson, T. R., & Zhang, J. (2005). A user-centered framework for redesigning healthcare interfaces. *Journal of Biomedical Informatics, 38*(1), 75–87. https://doi.org/10.1016/j.jbi.2004.11.005

Lai, Y.-S., Kankanhalli, A., & Ong, D. C. (2021). Human–AI collaboration in healthcare: A review and research agenda. *Proceedings of the 54th Hawaii International Conference on System Sciences.* https://doi.org/10.24251/hicss.2021.046

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Liu, Y., Han, S., & Zhang, W. (2024). Interactive explainability in clinical AI decision support: A mixed-method study. *Journal of Medical Internet Research, 26*, e58947. https://doi.org/10.2196/58947

Miller, T. (2023). Rethinking explainability: Adaptive transparency and human-centered AI. *AI & Society, 38*(2), 345–360. https://doi.org/10.1007/s00146-021-01342-0

Nazar, M., Alam, M. M., Yafi, E., & Mazliham, M. S. (2021). A systematic review of human–computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. *IEEE Access, 9,* 1–1. https://doi.org/10.1109/ACCESS.2021.3127881

Nazar, R., Khan, A., & Malik, M. (2021). Transparency in medical AI: Trust, adoption, and ethics. *Frontiers in Artificial Intelligence, 4*, 153316. https://doi.org/10.3389/frai.2021.715331

Nguyen, Q., Patel, A., & Green, M. (2023). Evaluating the impact of explainable AI on clinician trust and diagnostic accuracy. *Nature Medicine, 29*(8), 1760–1772. https://doi.org/10.1038/s41591-023-02571-4

Rajkomar, A., Chen, K., & Lungren, M. P. (2023). The next frontier of trustworthy AI in healthcare. *Nature Medicine, 29*(1), 37–45. https://doi.org/10.1038/s41591-022-02067-1

Rajpurkar, P., Chen, E., & Lungren, M. (2022). AI in medicine: Explainability, accountability, and trust. *The Lancet Digital Health, 4*(9), e646–e654. https://doi.org/10.1016/S2589-7500(22)00080-3

Sagona, A., Dai, T., Macis, M., & Darden, M. (2025). Trust in AI-assisted health systems and AI's trust in humans. *NPJ Health Systems, 2*(1). https://doi.org/10.1038/s44401-025-00016-5

Sagona, A., Yoon, J., & Blut, M. (2025). Trust in healthcare automation: Exploring AI transparency and reliability. *Frontiers in Digital Health, 7*, 112895. https://doi.org/10.3389/fdgth.2025.112895

Shortliffe, E. H., & Sepúlveda, M. J. (2021). Clinical decision support in the era of artificial intelligence. *JAMA, 325*(6), 512–523. https://doi.org/10.1001/jama.2020.25945

Smith, D., Patel, V., & Rahman, A. (2024). Explainable AI radiology systems improve interdisciplinary collaboration. *The Lancet Digital Health, 6*(4), e245–e257. https://doi.org/10.1016/S2589-7500(24)00052-9

Vasey, B., Clifton, D. A., & Collins, G. S. (2022). DECIDE-AI: Standards for clinical evaluation of healthcare AI. *Nature Medicine, 28*(8), 1507–1514. https://doi.org/10.1038/s41591-022-01985-7

Yoon, J., Blut, M., Wang, C., & Wunderlich, N. (2025). Human–robot collaboration and trust in healthcare AI systems. *Frontiers in Psychology, 16*, 133875. https://doi.org/10.3389/fpsyg.2025.133875

Yoon, J., Zajac, P., Pococke, L., Jicol, A., Clarke, C., O'Neill, E., Petrini, K., Lutteroth, C., & Jicol, C. (2025). The AI of the beholder: Experience of healthcare AI robots is shaped by user-centred factors, not their visual appearance. *[Journal name, volume(issue), page range]*. https://doi.org/xxxxxxxxx

Zhou, X., Li, C., & Lee, S. (2024). Cultural perceptions of AI transparency in global healthcare systems. *Frontiers in Public Health, 12*, 1398223. https://doi.org/10.3389/fpubh.2024.1398223