AI Privacy and Security in Healthcare: A Systematic Literature Review

Diane Dolezel, PhD[1], Karima Lalani, PhD[2], Valerie Watzlaf, PhD[3], Kerryn Butler-Henderson[4], Elise V.Z. Lambert[1], Mary Morton, PhD[5], Jamie Sand, EdD[6], David Gibbs, PhD[1], Susan Fenton, PhD[7]

[1] Health Informatics & Information Management Department, Texas State University, Round Rock, Texas, USA

[2] Health Informatics and Health Information Management Program, School of Public Health, University of Washington, Seattle, Washington, USA

[3] Department of Health Information Management, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

[4] Charles Sturt University, Wagga Wagga, New South Wales, Australia

[5] University of Mississippi Medical Center, Health Informatics and Information Management, Jackson, Mississippi, USA

[6] School of Public and Population Health, Boise State University, Boise, Idaho, USA

[7] McWilliams School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, Texas, USA

Abstract

Background: Artificial intelligence is expanding into telemedicine and telerehabilitation, yet significant privacy and security concerns persist. **Scope:** To synthesize empirical evidence on privacy and security approaches in health care, particularly those relevant to distributed home care. **Methodology:** A systematic review identified 80 studies (2019 to 2025), and Latent Dirichlet Allocation (LDA) topic modeling characterized the privacy and security themes. **Results:** Sixty-six studies addressed privacy, only seventeen addressed security, and three studies addressed both. LDA identified four themes: patient data privacy, federated learning for medical imaging, encrypted training and secure computation, and healthcare data governance. Most studies emphasized privacy-preserving approaches, like federated learning, encryption, and differential privacy. Almost half were conducted outside healthcare environments, limiting insight into real teleclinical and telerehabilitation workflow. **Conclusion:** Securing healthcare AI will require a multi-layered governance framework, broader global representation, and integration of privacy and security protections into routine clinical workflows.

Keywords: Artificial intelligence, Privacy, Security, Systematic review, Telerehabilitation

Artificial intelligence (AI) is increasingly integrated into clinical workflows, from prediction and image analysis to clinical decision support and remote care, which is expanding access but creating new exposure points for privacy and security. Privacy refers to a person's ability to determine when, how, and for what purposes their health information is collected, shared, and used (U.S. Department of Health & Human Services, 2023a). In healthcare, security is operationalized through administrative, technical, and physical safeguards that protect information systems and preserve confidentiality, integrity, and availability (U.S. Department of Health & Human Services, 2023b). Because AI systems rely on large, distributed, and continuously generated health data, these distinctions are important for governance and oversight. Despite a growing focus on AI innovation, privacy and security challenges specific to AI-enabled healthcare, especially in remote care contexts, remain fragmented and insufficiently explored.

AI Applications and Remote Health Care

In telehealth and telerehabilitation settings, AI extends care delivery beyond traditional clinical environments by supporting remote monitoring, clinical decision-making, documentation, and personalized care outside the hospital (Li, Li, Wei, & Li, 2024). AI applications include virtual assistants, wearable-based monitoring systems, predictive models of disease progression, personalized treatment recommendations, and AI-enabled clinical scribes (Chandrasekaran & Moustakas, 2025). As sensitive data flows through home networks, personal devices, and third-party platforms, privacy and security risks increase (Houser et al., 2023). Remote sensors and cloud-based systems further expand the attack surface for system-level vulnerabilities and coordination failures, highlighting the need for robust security and governance measures (Singh et al., 2025).

Privacy, Security, and Patient Trust in Remote Care

Successful adoption of AI-enabled telehealth depends on patient trust, particularly transparency in data collection and use (Chan et al., 2025; Chew & Achananuparp, 2022; L.Li et al., 2024). Although patients acknowledge AI's potential benefits, privacy concerns and limited understanding often reduce their willingness to share health data (Aggarwal et al., 2021; Davis et al., 2025; Khullar et al., 2022). For instance, while over half of U.S. respondents believed AI could improve care, most expressed privacy concerns, with similar patterns in Canada and the United Kingdom, where fears of re-identification further limited data sharing (Aggarwal et al., 2021; Davis et al., 2025; Khullar et al., 2022). Acceptance varies by application, with greater hesitation toward AI documentation tools but higher perceived privacy when using chatbots for sensitive topics (Chandrasekaran & Moustakas, 2025; Dellavelle et al., 2025). Prior reviews highlight gaps in usability, accountability, informed consent, and equity, underscoring the need for patient-centered co-design in AI governance (Moy et al., 2024).

Governance and Regulatory Challenges for Healthcare AI

In this review, governance refers to the policies, accountability structures, and operational controls that translate ethical principles into enforceable privacy and security practices. Despite regulatory efforts driven by patient concerns and organizational risk, the effectiveness of AI-enabled telehealth in addressing privacy and security challenges remains unclear. Chan et al. (2025) identified over 200 AI ethics frameworks but found limited real-world impact, particularly in translating principles into enforceable governance, security, and accountability mechanisms, leaving gaps in oversight for data misuse and AI-related harm.

Telehealth and telerehabilitation increases risk by extending data flows beyond clinical settings to personal devices, home networks, and third-party platforms. Regulatory oversight varies internationally: the American Telemedicine Association provides telehealth-specific guidance (American Telemedicine Association, 2025), whereas Software as a Medical Device is governed by region-specific frameworks, including the EU Medical Device Regulation, Australia's Therapeutic Goods Act, U.S. Food and Drug Administration guidance, and South Korea's Digital Medical Products Act (Dobson & Ruhl, 2025). International bodies such as the World Health Organization and the National Academy of Medicine have also proposed ethical frameworks that promote fairness and accountability (Aggarwal et al., 2021; Kelly et al., 2019; Tsoi, 2025).

Gaps in Literature and Rationale for Review

Most prior research has focused on AI adoption, public perceptions, and high-level governance frameworks, with comparatively less emphasis on privacy and security risks in AI-enabled healthcare. Studies highlight persistent mistrust, including concerns about unauthorized data collection from voice-activated devices (Chew & Achananuparp, 2022), and show that fewer than 25% of AI implementations report using privacy or security controls (Aravazhi et al., 2025). Although some reviews examine clinician cybersecurity duties or technical safeguards (Elendu et al., 2024; Shojaei et al., 2024), they do not address risks specific to AI-driven systems. These gaps point to the need for methods that can synthesize diverse evidence

and systematically identify recurring privacy and security issues. Combining Latent Dirichlet Allocation (LDA) with PRISMA provides a scalable, data-driven approach to detecting key themes and trends in this literature.

Research Questions

Guided by this combined systematic and data-driven approach, this review addresses the following research questions:

1. What are the key privacy and security themes in healthcare AI literature?

2. How can LDA enhance the identification of privacy and security themes in healthcare AI literature?

These findings are intended to inform policymakers, health informatics professionals, and AI developers seeking to deploy AI responsibly in telehealth and telerehabilitation settings.

Methods

The review process was organized in alignment with PRISMA guidance to ensure clear, comprehensive, and reproducible reporting (Page et al., 2021).

Eligibility Criteria

To be included, studies had to present original empirical findings, be published in English between 2018 and 2025, and focus on privacy or security issues within healthcare-related AI. Publications that did not provide primary data, such as systematic reviews, meta-analyses, dissertations, theses, or editorial and opinion pieces, were excluded to ensure the review synthesized only firsthand evidence. Non-English studies were also omitted to support consistency in coding and interpretation.

Information Sources and Search Strategy

A comprehensive search strategy was developed to identify relevant studies across seven databases: CINAHL, EBSCO, Ovid, PubMed, ScienceDirect, Scopus, and Web of Science. The year 2018 was selected as the start date, as it marks a period in which the adoption of AI in healthcare began to accelerate substantially (Jackson & Hu, 2019). Search terms were designed to capture literature related to healthcare AI privacy and security issues, using this Boolean string:

[artificial intelligence OR automation OR augmented] AND

[privacy OR security] AND

[health OR clinical OR clinician OR medical OR medicine OR nurse OR nursing OR patient].

Selection Process

Two members of the nine-person review team (DD, KL, EL, VW, JS, KBH, MM, DG, JM) independently screened each title and abstract using the predetermined inclusion and exclusion criteria, resolving differing assessments through discussion. Full-text evaluation was then conducted by reviewer pairs drawn from a group of seven reviewers (DD, KL, JS, MM, VW, EL, DG), with any unresolved differences referred to a third reviewer for final judgment. Study details, such as privacy or security focus, target population, intervention characteristics, outcomes, and major findings, were recorded using a standardized extraction template within Covidence® (Veritas Health Innovation, n.d.).

Data Extraction

All identified records were imported into Covidence®, an online platform for systematic review management, where duplicates were removed. The data extraction form was pilot tested by two reviewers before full extraction to ensure consistency across reviewers.

Thematic Analysis

Eighty studies met the eligibility criteria and were exported as Word files for processing in R, where text documents were prepared using the *topicmodels* package. LDA was then applied to uncover dominant themes by examining patterns of word co-occurrence across the text body (Blei et al., 2003; Feller et al., 2018). Preprocessing steps involved routine text-cleaning procedures, including the removal of punctuation and stop words. Multiple model configurations were tested, and a four-topic solution was selected because it offered the most coherent and interpretable structure. Two reviewers independently reviewed the LDA outputs and refined the topic labels, organizing them into broader thematic categories based on salient terms and conceptual alignment. The use of LDA in this context is consistent with prior healthcare research that employs topic modeling to analyze clinical narratives and other unstructured text sources (Danler et al., 2024; Sun et al., 2024). A formal risk-of-bias assessment was not performed, given the substantial heterogeneity in study designs and the exploratory aims of the analysis.

Results

Study Selection

Figure 1 summarizes the PRISMA study selection workflow. The database search yielded 1,688 records, of which 232 duplicates were removed before screening. The remaining 1,456 titles and abstracts were assessed, of which 1,194 did not meet the inclusion criteria. A total of 262 full-text articles were reviewed, and 182 were excluded for reasons such as absence of an AI component (n = 37), unsuitable study design (n = 32), a focus limited to cybersecurity rather than privacy or security (n = 30), an ineligible study setting (n = 21), or lack of empirical data (n = 14). The other 48 studies were excluded due to many different reasons as can be seen in Figure 1. Ultimately, 80 studies met all criteria and were included in the final analysis.

Characteristics of Included Studies

Table 1 provides an overview of the study characteristics, including design type, study setting, and publication year for all 80 included studies. The studies varied by design, setting, and publication year.

Study designs were grouped according to the primary methodological approach described by each article. Experimental research was the most common (n = 45, 56.2%), followed by model development (n = 12, 15.0%). Case studies and pilot investigations were less frequent (each n = 4, 5.0%), as were cohort and cross-sectional designs (each n = 3, 3.8%). A small number of studies employed alternative approaches such as algorithmic development, qualitative methods, mixed-methods designs, or exploratory formats, and one study did not specify a design.

Study settings varied widely, with nearly half of the research conducted outside of traditional healthcare environments (n = 37, 46.2%). Hospital studies accounted for 17.5% (n = 14). Thirteen articles (16.2%) did not report on their study setting.

Figure 1

PRISMA Study Selection Workflow

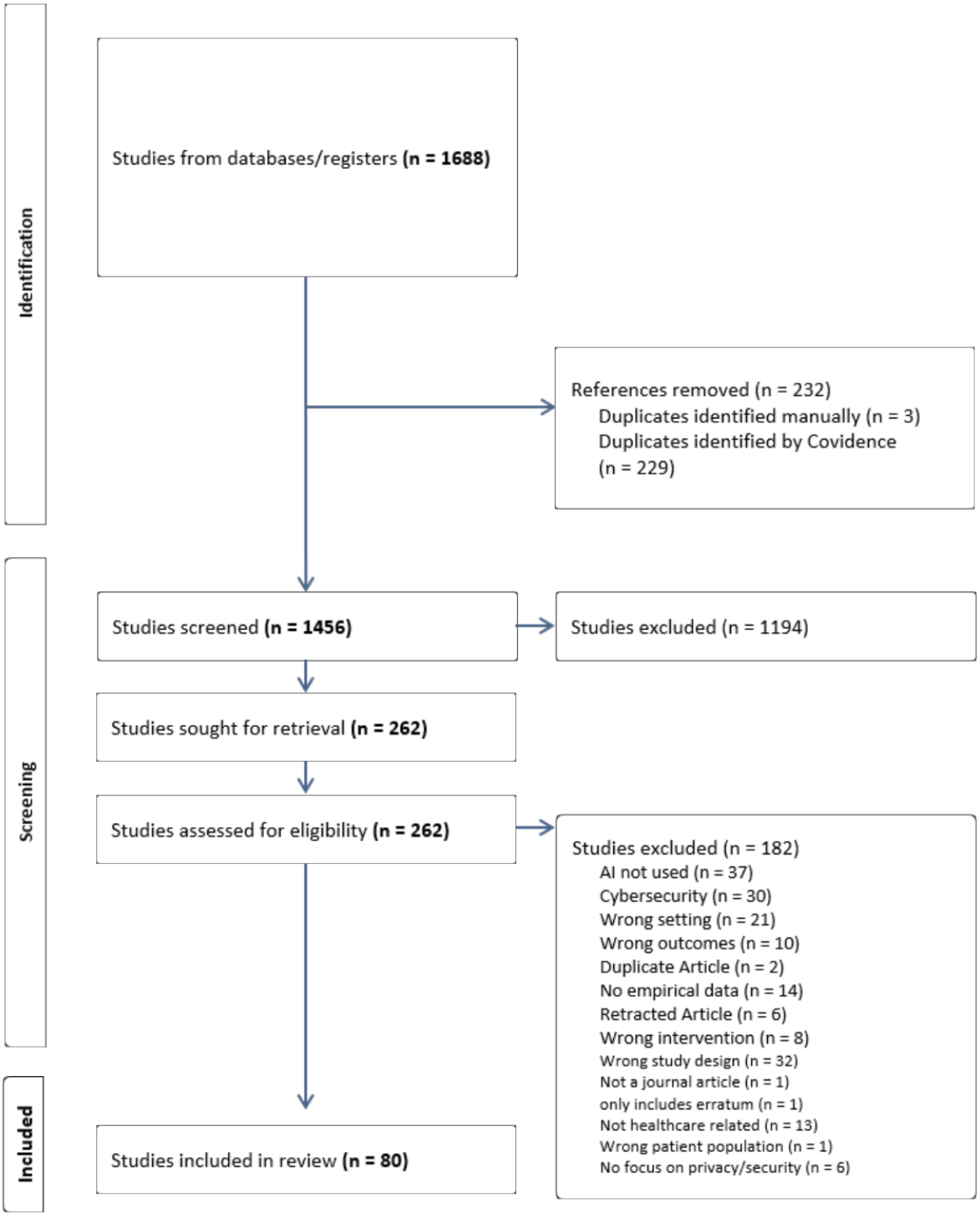


Table 1

Descriptive Characteristics of Included Studies (n=80)

Characteristics	Number	Percent (%)
Study Design		
Experimental	45	56.2
Model Development	12	15.0
Case Study	4	5.0
Pilot Study	4	5.0
Cohort	3	3.8
Cross-sectional	3	3.8
Algorithmic Development	2	2.5
Qualitative	2	2.5
Case-Control	1	1.2
Exploratory	1	1.2
Infrastructure Evaluation	1	1.2
Mixed Methods	1	1.2
Not Available	1	1.2
Study Site		
Non-healthcare Setting	37	46.3
Other	16	20.0
Hospital	14	17.5
Not Available	13	16.3

Note. "Not Available" reflects incomplete reporting by the authors.

Distribution of Included Studies by Year

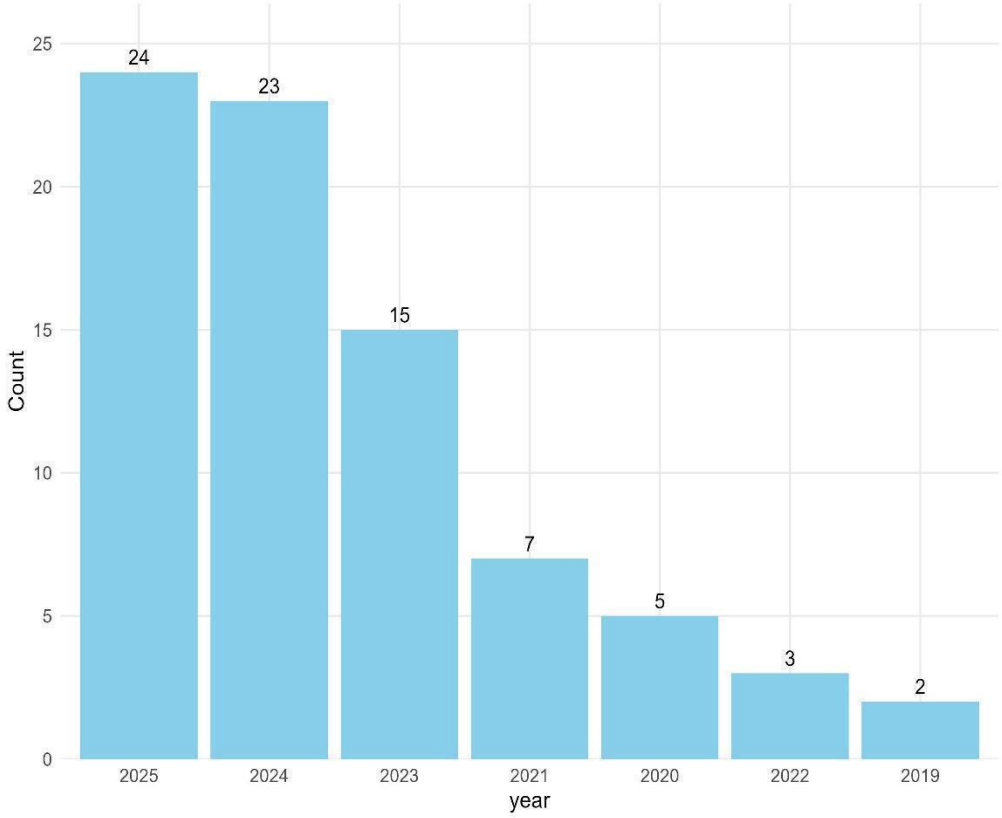
Figure 2 displays the distribution of studies published by year from 2019 to 2025. Study publications increased over time, with peaks in 2024 (n=23) and 2025 (n=24). Fewer studies were published before 2023, particularly in 2019 (n = 2) and 2022 (n = 3). One study did not report a publication year.

Descriptive Summaries of Journals Reviewed

The Appendix (Table A1) presents a detailed summary of the 80 studies examining privacy and security content from 2019 to 2025. Although the search strategy included studies from 2018 onward, no studies published in 2018 met the inclusion criteria. Among the 80 included studies, 66 (82.5%) addressed privacy and 17 (21.25%) explored security, with three of these studies addressing both topics.

Figure 2

Distribution of Studies by Publication Year, 2019–2025



Geographic Distribution of Included Studies

Table 2 summarizes the geographic distribution of the included studies. The table reports the number of studies conducted in each country, noting that some projects involved multiple international sites; therefore, frequencies are presented as counts rather than percentages. Research activity was concentrated primarily in Asia and the Middle East. China contributed the largest number of studies (n = 25), followed by India (n = 16) and Saudi Arabia (n = 10). The United States accounted for six studies, while Canada and Germany each contributed five, and South Korea contributed four. Several additional countries, including Bangladesh, Finland, Jordan, Malaysia, the Netherlands, Pakistan, and Taiwan, were represented by two studies each. The remaining countries, Algeria, Australia, Austria, Belgium, Hong Kong, Oman, Portugal, Romania, Slovakia, South Africa, Turkey, the United Kingdom, and Yemen, each appeared in one study.

Table 2

Geographic Distribution by Frequency

Country	Number
China	25
India	16
Saudi Arabia	10
United States	6

Canada	5
Germany	5
South Korea	4
Bangladesh	2
Finland	2
Jordan	2
Malaysia	2
Netherlands	2
Pakistan	2
Taiwan	2
Algeria	1
Australia	1
Austria	1
Belgium	1
Hong Kong	1
Oman	1
Portugal	1
Romania	1
Slovakia	1
South Africa	1
United Arab Emirates	2
Turkey	1
United Kingdom	1
Yemen	1

LDA Findings

After summarizing study characteristics and publication patterns, we applied LDA to uncover underlying themes in the literature. The LDA topic modeling identified four themes, each characterized by its most frequent keywords:

1. Patient Data Privacy in ML Models (keywords: data, privacy, learning, model, health, dataset).
2. Federated Learning for Medical Imaging Privacy and Performance (keywords: federated, privacy, training, performance, medical, image).
3. Encrypted Medical Imaging and Secure Training (keywords: encryption, privacy, learning, training); and
4. Healthcare Data Privacy and System Governance (keywords: healthcare, system, security, information).

Federated learning is discussed in multiple thematic areas, including decentralized training, privacy protection, and secure collaboration frameworks.

Thematic Results

Theme 1: Patient Data Privacy in ML Models

The first theme centers on the protection of patient information within healthcare AI systems, a dominant concern across the included studies. Terms such as *privacy*, *data*, *health*, *information*, and *learning* frequently appeared, reflecting the strong emphasis on safeguarding sensitive clinical details. A range of privacy-preserving techniques were described, including federated learning to support decentralized model development for applications such as disease prediction, mental-health assessment, EHR exchange, and medical-imaging analysis (Choudhury et al., 2025; Dubey et al., 2025; Lee et al., 2024; Om Kumar et al., 2023; Shao et al., 2020; Xu et al., 2022). Additional strategies, such as differential privacy, blockchain infrastructures, homomorphic encryption, k-anonymity, and entropy-based methods, were also used to minimize re-identification risks (Alabdulwahab et al., 2024; Benouis et al., 2024; Hennebelle et al., 2024; Kokin et al., 2024; Sangaiah et al., 2024; S.Wang et al., 2025).

Even with these safeguards, literature highlights several vulnerabilities that remain unresolved. Studies have shown that word-embedding models can inadvertently expose patient details (Abdalla et al., 2020), large language models may reveal protected information under certain conditions (Chuang et al., 2025), and behavioral-biometric data can still enable individual re-identification (Deepak et al., 2024; Vignesh et al., 2025). To mitigate these risks, researchers increasingly advocate for local model deployment or the use of open-source, privacy-preserving tools that limit or avoid external data transfer (Nowak et al., 2025; Seyfizadeh et al., 2024; Y.Zhang et al., 2025).

Theme 2: Federated Learning for Medical Imaging Privacy and Performance

The second theme emphasizes the prominent role of federated learning (FL) in supporting privacy-preserving development of medical-imaging AI models. The frequent appearance of keywords such as *federated*, *privacy*, *medical*, and *image* reflects how strongly FL is associated with decentralized analytics across healthcare institutions. By enabling algorithms to be trained locally at each site while sharing only model parameters, FL provides a practical solution for multi-institutional collaboration in contexts where direct data sharing is restricted by ethical, legal, or regulatory requirements. This approach has been used across a wide range of imaging applications, including skin cancer classification, tumor characterization, ophthalmic disease detection, polyp segmentation, sleep-stage recognition, ECG analysis, and COVID-19 assessment (Ain et al., 2023; Baumgartner et al., 2023; Butt et al., 2023; Hasan & Rahman, 2025; Li et al., 2025; Yahiaoui et al., 2024; Yan et al., 2023).

A consistent challenge across these studies is the substantial heterogeneity in data distributions between institutions, which can degrade model performance when training is performed locally. Researchers have addressed this issue using strategies such as domain adaptation, transfer learning, and personalized aggregation mechanisms designed to adjust model parameters to site-specific characteristics (Bangare et al., 2024; Gulati et al., 2025; Haripriya et al., 2025; Tang et al., 2024). Many FL implementations also introduce additional privacy safeguards, including differential privacy, synthetic data augmentation, and noise-based perturbation, to reduce the risk of re-identification and to support learning from scarce or imbalanced datasets (Hasan & Rahman, 2025; He et al., 2024; Iqbal et al., 2025; Yahiaoui et al., 2024; Yang et al., 2025; L.Zhang et al., 2025).

However, the diversity of FL architecture and analytic pipelines used across studies makes direct comparison difficult. Differences in model aggregation algorithms, encryption strategies, communication protocols, client-side configurations, and privacy-preserving add-ons contribute to inconsistent reporting and hinder the ability to standardize findings across the literature (Chen et al., 2024; J.Wang, 2025; Q.Wang et al., 2025).

Theme 3: Encrypted Medical Imaging and Secure Training

The third theme captures research that applies encryption-based and cryptographic techniques to protect medical images during AI model development. The frequent appearance of terms such as *medical*, *imaging*, *encryption*, and *training* reflects the emphasis on securing data throughout the learning pipeline. Several studies demonstrate that models can be trained on encrypted inputs, using approaches such as homomorphic encryption and secure multi-party computation, while still achieving

accuracy levels comparable to those trained on unencrypted data, although typically with increased computational demands (Son et al., 2021; Vizitiu et al., 2019). Work spanning cardiac monitoring, breast cancer diagnostics, and other imaging domains shows that strong encryption mechanisms can effectively protect patient information without severely compromising downstream model performance (Bai et al., 2025; Cao et al., 2023; L.Chen et al., 2024).

Another area of research focused on creating secure collaborative environments for AI development without sharing raw patient data. These efforts incorporate blockchain-enabled platforms, federated or distributed learning frameworks, and other privacy-preserving infrastructures designed to maintain data confidentiality while still supporting cross-institutional modeling efforts (Barnawi et al., 2024; Guan et al., 2023; Kaissis et al., 2021; Zerka et al., 2020). Collectively, these studies point to a growing interest in embedding cryptographic safeguards directly into medical AI workflows to reduce exposure risks and strengthen trust in distributed analytic systems.

Theme 4: Healthcare Data Privacy and System Governance

The fourth theme captures the broader governance and organizational structures necessary to support privacy and security in healthcare AI systems. The recurring presence of terms such as *healthcare*, *system*, *security*, and *information* reflects the emphasis on institutional oversight rather than solely technical solutions. Across the included studies, researchers described governance approaches that integrate technical safeguards with organizational policies, for example, the use of privacy-impact assessments, audit trails, role-based access controls, encryption protocols, and secure logging to align AI operations with existing IT infrastructures (Danler et al., 2024; Sun et al., 2024).

A recurring challenge involves ensuring compliance with regulatory frameworks when AI systems span multiple institutions or areas. Studies noted that meeting the requirements of the Health Insurance Portability and Accountability Act (HIPAA), General Data Protection Regulation (GDPR), and region-specific privacy and security laws becomes increasingly complex as data moves across organizational boundaries (Dubey et al., 2025; Lee et al., 2024). Danler et al. (2024) specifically highlighted organizational strategies for bringing AI workflows into alignment with hospital IT policies, such as enforcing role-based permissions, implementing encryption standards, and maintaining tamper-resistant logs. Additional research also examined ethical and legal considerations associated with cross-site data flows, reinforcing that regulatory obligations are more difficult to satisfy when AI solutions are deployed in distributed environments (Dubey et al., 2025; Lee et al., 2024).

Beyond policy compliance, several studies explored how AI technologies operate within routine clinical information environments. These investigations discussed the integration of AI into electronic health records and health information exchange systems, alongside mechanisms such as federated learning, system-level monitoring, and well-defined data-stewardship responsibilities. Auditing tools and risk-management practices were also described as important mechanisms for preventing unintended disclosure and supporting accountable decision-making within AI-augmented processes (Om Kumar et al., 2023; Seyfizadeh et al., 2024). Collectively, this body of work underscores the increasing recognition that strong governance frameworks are essential complements to technical privacy controls, particularly in telehealth and telerehabilitation contexts where data often flows outside traditional institutional boundaries.

Discussion

This systematic review brings together recent research addressing privacy and security concerns related to the use of AI in healthcare. The study literature is largely dominated by experimental studies, indicating a primary emphasis on the design and technical evaluation of AI-based solutions, with comparatively limited engagement in theory-driven or exploratory inquiry. Across studies, privacy protections receive more sustained attention than security-focused measures, including intrusion detection, threat surveillance, and resilience to adversarial attacks. Privacy and security are often discussed in tandem, which may conceal system-level weaknesses that are unique to AI-enabled clinical technologies. As a result, the body of evidence reflects a stronger orientation toward protecting sensitive information than toward safeguarding the operational integrity of AI systems deployed in clinical contexts.

These challenges are amplified in telehealth and telerehabilitation settings, where AI-supported services commonly function within decentralized, patient-managed environments rather than within centrally controlled institutional infrastructures. Remote monitoring technologies, wearable devices, and video-based care platforms generate continuous streams of sensitive data beyond traditional clinical boundaries. In home-based care scenarios, informed consent processes may become more

difficult to implement effectively, particularly when patients are expected to understand complex privacy-preserving methods such as federated learning or differential privacy despite varying levels of technical familiarity.

A variety of technical strategies, including federated learning, differential privacy, blockchain-based architecture, encryption methods, and locally deployed language models, have been proposed to mitigate these risks. Although these approaches are well suited to decentralized and remote-care contexts, their application and performance within routine telehealth and telerehabilitation workflows remain inconsistently examined. Taken together, the findings point to persistent gaps in the practical integration of privacy and security mechanisms within remote-care ecosystems, offering important context for the detailed results, comparisons with prior work, and implications discussed in the sections that follow.

Main Findings

Privacy and security concerns are particularly pronounced in telehealth and telerehabilitation settings, where continuous data capture, home-based monitoring technologies, and reliance on third-party platforms can increase exposure risks while reducing direct clinical oversight. Of the 80 studies included in this review, 66 focused on privacy-related issues, whereas only 17 explicitly examined security concerns, underscoring an imbalance in the current research landscape. Federated learning emerges as a frequently investigated strategy for enabling cross-institutional collaboration without the need to centralize patient data (Ain et al., 2023; Butt et al., 2023; Choudhury et al., 2025; Yan et al., 2023). Many studies further combine federated learning with techniques such as differential privacy and secure aggregation to limit the risk of sensitive information being inferred from model updates (He et al., 2024; Yang et al., 2025; L.Zhang et al., 2025).

Encryption remains a core strategy for securing health information in transit and at rest, while advanced methods, including homomorphic encryption and secure multi-party computation, support computation without direct exposure of raw data (P.-Y Chen et al., 2024; Kumar et al., 2023). Blockchain-based systems have also been explored as tools for enhancing data integrity, managing access permissions, and improving auditability within distributed clinical environments. Taken together, these findings show that most research continues to prioritize methods for protecting data confidentiality, whereas far fewer studies evaluate the operational security of AI systems once deployed in clinical settings.

Comparison with Previous Research

Earlier scholarship consistently described privacy and security as foundational concerns in the use of AI within healthcare, frequently noting ethical challenges, regulatory uncertainty, public trust, and inconsistent adoption of available safeguards (Aggarwal et al., 2021; Chan et al., 2025; Chew & Achananuparp, 2022; Kelly et al., 2019; Khuller et al., 2022; Y.Li, et al., 2024; Tsoi, 2025). Across these studies, patient trust has been closely associated with transparent data practices and the presence of strong privacy protection. In contrast, the findings of the present review suggest a growing emphasis on system-level and architectural approaches that incorporate privacy and security directly into the design of AI-enabled healthcare technologies.

More recent work frames privacy engineering and governance as integral operational components, linking them to health-information exchange processes and model-performance considerations (Ragab et al., 2024; L.Zhang et al., 2025). Federated learning, once discussed primarily as a conceptual tool (Shojaei et al., 2024), is now more commonly implemented in multi-layered defense strategies aimed at protecting data while facilitating data analysis (Ain et al., 2023; Baumgartner et al., 2023; Butt et al., 2023; Hasan & Rahman, 2025; Li et al., 2025; Yahiaoui et al., 2024; Yan et al., 2023). Governance models have likewise evolved from abstract principles toward operational mechanisms, including access control systems, AI-enabled monitoring tools, and blockchain-supported electronic health records (Anita et al., 2025; Hennebelle et al., 2024; Ibrahim et al., 2024; Supraja et al., 2025). Taken together, this progression reflects an increasingly integrated approach in which privacy and security are addressed across the entire AI lifecycle to support scalable, trustworthy innovation in healthcare.

Implications for Future Practice and Research

The findings of this systematic review indicate that privacy and security weaknesses persist in AI-enabled healthcare systems, even when data have been de-identified. This reality highlights the importance of embedding privacy-preserving strategies, such as federated learning, differential privacy, encryption, and secure computation, directly into clinical AI

workflows. These protections are especially crucial in telehealth and remote-monitoring environments, where data are generated, transmitted, and processed outside the safeguards typically present in controlled clinical settings.

However, technical safeguards on their own are insufficient. Variability in data quality, institutional processes, and patient populations across decentralized systems can introduce bias, compromise model stability, and degrade diagnostic performance, particularly within federated learning architectures. Ensuring safe and scalable implementation will require healthcare organizations to establish coordinated governance frameworks that include standardized data-quality controls, clear procedures for model updating and aggregation, continuous performance auditing, and compliance with regulatory requirements such as HIPAA and GDPR.

Although encrypted training pipelines and secure imaging processes can further reduce exposure risk, wide-scale adoption is constrained by computational overhead and infrastructure limitations. Future work should prioritize real-world evaluations of these privacy-preserving methods, including assessments of their impact on model accuracy and reliability, as well as investigations into patient understanding of AI-related data protections to strengthen informed-consent practices.

Study Limitations

This review has several limitations. First, the geographic representation was uneven, with most studies conducted in China, India, and Saudi Arabia, and few from the U.S, which may limit generalizability across different regulatory and cultural contexts. Second, many studies used experimental designs in non-healthcare settings, reducing insight into practical applicability. Third, the diversity of AI models (e.g., Convolutional Neural Network (CNN), Generative Adversarial Network (GAN), etc.) and privacy tools (e.g., differential privacy, blockchain) made direct comparisons problematic. Fourth, limited reporting on security measures may bias findings. Fifth, there was incomplete reporting of several study variables, which may introduce a reporting bias. Finally, limiting studies to English may introduce language bias and may exclude relevant research.

Conclusions

This systematic review highlighted four key privacy and security themes in healthcare AI, showing a strong focus on privacy techniques like federated learning but limited attention to security measures and real-world use. Strengthening trustworthy AI will require multi-layered governance, consistent standards, broader geographic representation, and closer alignment between system design and health data practices.

Corresponding Author

Valerie Watzlaf, PhD

valgeo@pitt.edu

References

Aanjankumar, S., Muchahari, M.K., Urooj, S., et al. (2025). Enhanced consumer healthcare data protection through AI-driven TinyML and privacy-preserving techniques. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3573076

Abdalla, M., Abdalla, M., Hirst, G., & Rudzicz, F. (2020). Exploring the privacy-preserving properties of word embeddings: Algorithmic validation study. *Journal of Medical Internet Research*, *22*(7), e18055. https://doi.org/10.2196/18055

Aggarwal, R., Farag, S., Martin, G., Ashrafian, H., & Darzi, A. (2021). Patient perceptions on data sharing and applying artificial intelligence to health care data: Cross-sectional survey. *Journal of Medical Internet Research, 23*(8), e26162. https://doi.org/10.2196/26162

Ain, Q. U., Khan, M. A., Yaqoob, M. M., Khattak, U. F., Sajid, Z., Khan, M. I., & Al-Rasheed, A. (2023). Privacy-aware collaborative learning for skin cancer prediction. *Diagnostics (Basel)*, *13*(13). https://doi.org/10.3390/diagnostics13132264

Akmeşe, M. F. (2024). Data privacy-aware machine learning approach in pancreatic cancer diagnosis. *BMC Medical Informatics and Decision Making, 24,* 248. https://doi.org/10.1186/s12911-024-02657-2

Alabdulwahab, S., Kim, Y. T., & Son, Y. (2024). Privacy-preserving synthetic data generation method for IoT-sensor network IDS using CTGAN. *Sensors (Basel)*, *24*(22). https://doi.org/10.3390/s24227389

Almaiah, M. A., Yelisetti, S., Arya, L., Babu Christopher, N. K., Kaliappan, K., Vellaisamy, P., Hajjej, F., & Alkdour, T. (2023). A novel approach for improving the security of IoT–medical data systems using an enhanced dynamic Bayesian network. *Electronics*, *12*(20). https://doi.org/10.3390/electronics12204316

Almufareh, M. F. (2024). An edge computing-based factor-aware novel framework for early detection and classification of melanoma disease through a customized VGG16 architecture with privacy preservation and real-time analysis. *IEEE Access*, *12*, 113580–113596. https://doi.org/10.1109/access.2024.3444050

American Telemedicine Association. (2025). *Policy principles on artificial intelligence*. American Telemedicine Association. https://info.americantelemed.org/hubfs/Files/Policy%20Docs_letters%2c%20RFI%2c%20etc./ATA%20AI%20Policy%20Principles_2025.pdf

Amoon, M., Altameem, T., & Altameem, A. (2020). IoT sensor assisted security and quality analysis for healthcare data sets using artificial intelligent-based heuristic health management system. *Measurement, 161*, 107861. https://doi.org/10.1016/j.measurement.2020.107861

Anita, C.S., Ananthajothi, K., & Joselin Jeya Sheela, J. (2025). A joint multimodal user authentication-based privacy preservation with disease prediction framework in modern healthcare system using multi-scale cross attention-based ResNet. *Computer Methods and Programs in Biomedicine, 270*, 108928. https://doi.org/10.1016/j.cmpb.2025.108928

Arafat, Y., Siam, A. S. M., Chowdhury, M. M., Hasan, M. M., Jobayer, S. H., Shatabda, S., Islam, S., & Mukta, S. (2025). Decentralized medical image sharing: A blockchain based approach with subject sensitive hashing for enhanced privacy and integrity. *IET Blockchain*, *5*(1). https://doi.org/10.1049/blc2.70009

Arasteh, S.T., Lotfinia, M., Nolte, T., Sahn, M.-J., Isfort, P., Kuhl, C., Nebelung, S., Kaissis, G., & Truhn, D. (2024). Securing collaborative medical AI by using differential privacy: Domain transfer for classification of chest radiographs. *Radiology: Artificial Intelligence, 6*(1), e230212. https://doi.org/10.1148/ryai.230212

Aravazhi, P. S., Gunasekaran, P., Benjamin, N. Z. Y., Thai, A., Chandrasekar, K. K., Kolanu, N. D., & Inban, P. (2025). The integration of artificial intelligence into clinical medicine: Trends, challenges, and future directions. *Disease-A-Month, 71*, 101882. https://doi.org/10.1016/j.disamonth.2025.101882

Bai, Y., Zhao, H., Shi, X., & Chen, L. (2025). Towards practical and privacy-preserving CNN inference service for cloud-based medical imaging analysis: A homomorphic encryption-based approach. *Computer Methods and Programs in Biomedicine*, *261*, 108599. https://doi.org/10.1016/j.cmpb.2025.108599

Bangare, J. L., Sable, N. P., Mahalle, P. N., & Shinde, G. R. (2024). A federated learning approach for non-co-located datasets: Enhancing data governance and privacy. *Frontiers in Health Informatics*, *13*(2). https://healthinformaticsjournal.com/index.php/IJMI/article/view/8

Barnawi, A., Chhikara, P., Tekchandani, R., Kumar, N., & Alzahrani, B. (2024). A differentially privacy assisted federated learning scheme to preserve data privacy for IoMT Applications. *IEEE Transactions on Network and Service Management*, *21*(4), 4686–4700. https://doi.org/10.1109/tnsm.2024.3393969

Baumgartner, M., Veeranki, S. P. K., Hayn, D., & Schreier, G. (2023). Introduction and comparison of novel decentral learning schemes with multiple data pools for privacy-preserving ECG classification. *Journal of Healthcare Informatics Research*, *7*(3), 291–312. https://doi.org/10.1007/s41666-023-00142-5

Bavli, I., Ho, A., Mahal, R., & McKeown, M. J. (2024). Ethical concerns around privacy and data security in AI health monitoring for Parkinson's disease: Insights from patients, family members, and healthcare professionals. *AI & Society*, *40*(1), 155–165. https://doi.org/10.1007/s00146-023-01843-6

Benouis, M., Andre, E., & Can, Y. S. (2024). Balancing between privacy and utility for affect recognition using multitask learning in differential privacy-added federated learning settings: Quantitative study. *JMIR Mental Health*, *11*, e60003. https://doi.org/10.2196/60003

Bergen, R. V., Rajotte, J. F., Yousefirizi, F., Rahmim, A., & Ng, R. T. (2024). Assessing privacy leakage in synthetic 3-D PET imaging using transversal GAN. *Computer Methods and Programs in Biomedicine*, *243*, 107910. https://doi.org/10.1016/j.cmpb.2023.107910

Bharath, K. N., Babu, K. S., & Ravi, V. (2024). Security of magnetic resonance medical images using region-based lossless image compression in healthcare information systems. *The Open Public Health Journal*, *17*(1). https://doi.org/10.2174/0118749445326668240801105910

Bi, H., Liu, J., & Kato, N. (2022). Deep learning-based privacy preservation and data analytics for IoT enabled healthcare. *IEEE Transactions on Industrial Informatics*, *18*(7), 4798–4807. https://doi.org/10.1109/tii.2021.3117285

Butt, M., Tariq, N., Ashraf, M., Alsagri, H. S., Moqurrab, S. A., Alhakbani, H. A. A., & Alduraywish, Y. A. (2023). A fog-based privacy-preserving federated learning system for smart healthcare applications. *Electronics*, *12*(19). https://doi.org/10.3390/electronics12194074

Chan, A., Rahimi-Ardabili, H., Rogers, W. A., & Coiera, E. (2025). The real-world impact of artificial intelligence ethics frameworks across a decade in healthcare: A scoping review. *Journal of the American Medical Informatics Association, 32*(11), 1767–1777. https://doi.org/10.1093/jamia/ocaf167

Chandrasekaran, R., & Moustakas, E. (2025). Patient attitudes toward ambient artificial intelligence scribes in clinical care: Insights from a cross-sectional study. *Journal of the American Medical Informatics Association*. Advance online publication. https://doi.org/10.1093/jamia/ocaf218

Chen, L., Song, L., Feng, H., Zeru, R. T., Chai, S., & Zhu, E. (2024). Privacy-SF: An encoding-based privacy-preserving segmentation framework for medical images. *Image and Vision Computing*, *151*. https://doi.org/10.1016/j.imavis.2024.105246

Chen, P.-Y., Cheng, Y.-C., Zhong, Z.-H., Zhang, F.-Z., Pai, N.-S., Li, C.-M., & Lin, C.-H. (2024). Information security and artificial intelligence–assisted diagnosis in an Internet of Medical Thing system (IoMTS). *IEEE Access*, *12*, 9757–9775. https://doi.org/10.1109/access.2024.3351373

Chew, H. S. J., & Achananuparp, P. (2022). Perceptions and needs of artificial intelligence in health care to increase adoption: Scoping review. *Journal of Medical Internet Research, 24*(1), e32939. https://doi.org/10.2196/32939

Choudhury, A., Volmer, L., Martin, F., Fijten, R., Wee, L., Dekker, A., & Soest, J. V. (2025). Advancing privacy-preserving health care analytics and implementation of the Personal Health Train: Federated deep learning study. *JMIR AI*, *4*, e60847. https://doi.org/10.2196/60847

Chuang, Y. S., Sarkar, A. R., Hsu, Y. C., Mohammed, N., & Jiang, X. (2025). Robust privacy amidst innovation with large language models through a critical assessment of the risks. *Journal of the American Medical Informatics Association*, *32*(5), 885–892. https://doi.org/10.1093/jamia/ocaf037

Davis, V. H., Qiang, J. R., Adekoya MacCarthy, I., Howse, D., Seshie, A. Z., Kosowan, L., Delahunty-Pike, A., Abaga, E., Cooney, J., Robinson, M., Senior, D., Zsager, A., Aubrey-Bassler, K., Irwin, M., Jackson, L. A., Katz, A., Marshall, E. G., Muhajarine, N., Neudorf, C., Garies, S., & Pinto, A. D. (2025). Perspectives on using artificial intelligence to derive social determinants of health data from medical records in Canada: Large multijurisdictional qualitative study. *Journal of Medical Internet Research, 27*, e52244. https://doi.org/10.2196/52244

Deepak, G., Sharma, P., Jayachitra, S., Chepur, J., Srihari, T., Judge, T. (2024). Privacy- preserving deep learning approaches for effective utilization of wearable health data. *Measurement: Sensors,* 33,101238. https://doi.org/10.1016/j.measen.2024.101238

Dellavalle, N. S., Ellis, J. R., Moore, A. A., Akerson, M., Andazola, M., Campbell, E. G., & DeCamp, M. (2025). What patients want from healthcare chatbots: Insights from a mixed-methods study. *Journal of the American Medical Informatics Association, 32*(11), 1735–1745. https://doi.org/10.1093/jamia/ocaf164

Dobson, S. J., & Ruhl, K. R. (2025, June 26). *From telemedicine to AI: The complex regulation of digital healthcare*. Ashurst. https://www.ashurst.com/en/insights/from-telemedicine-to-ai-the-complex-regulation-of-digital-healthcare/

Dubey, P., Dubey, P., Bokoro, P. (2025) Federated learning for privacy-enhanced mental health prediction with multimodal data integration. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization* https://doi.org/10.1080/21681163.2025.2509672

Dubey, S., Verma, D. (2022) Fuzzy logic based intelligent data sensitive security model for big data in healthcare. *International Journal of Electronics and Telecommunications, 68*(2), pp.245-250. https://doi.org/10.24425/ijet.2022.139874

Elendu, C., Omeludike, E. K., Oloyede, P. O., Obidigbo, B. T., & Omeludike, J. C. (2024). Legal implications for clinicians in cybersecurity incidents: A review. *Medicine, 103*(39): e39887. https://doi.org/10.1097/MD.0000000000039887

Ellis, J., Dellavalle, N., Hamer, M., Akerson, M., Andazola, M., Moore, A., Campbell, E., & DeCamp, M. (2025). The halo effect: Perceptions of information privacy among healthcare chatbot users. *Journal of the American Geriatrics Society, 73*(5), 1472–1483. https://doi.org/10.1111/jgs.19393

Ferdowski, M., Hasan, M., Habib, W. (2024) Responsible AI for cardiovascular disease detection: Towards a privacy-preserving and interpretable model. *Computer Methods and Programs in Biomedicine, 254* https://doi.org/10.1016/j.cmpb.2024.108289

Gong, X., Gao, J., Sun, S., Zhong, Z., Shi, Y., Zeng, H., & Yang, K. (2025). Adaptive compressed-based privacy-preserving large language model for sensitive healthcare. *IEEE Journal of Biomedical and Health Informatics.* https://doi.org/10.1109/JBHI.2025.3558935

Guan, X., Yu, G., Zhang, W., Wen, R., Wei, R., Jiao, S., Zhao, Q., Lou, Z., Hao, L., Liu, E., Gao, X., Wang, G., Zhang,W., & Wang, X. (2023) An easy-to-use artificial intelligence preoperative lymph node metastasis predictor (LN-MASTER) in rectal cancer based on a privacy-preserving computing platform: multicenter retrospective cohort study. *International Journal of Surgery, 109*(3), 255-265. https://dx.doi.org/10.1097/JS9.0000000000000067

Gulati, S., Guleria, K., & Goyal, N. (2025). Privacy-preserving and collaborative federated learning model for the detection of ocular diseases. *International Journal of Mathematical, Engineering and Management Sciences, 10* (1), 218-248, https://doi.org/10.33889/IJMEMS.2025.10.1.013

Guo, C., Jia, J., Choo, K., & Jie, Y. (2020). Privacy-preserving image search (PPIS): Secure classification and searching using convolutional neural network over large-scale encrypted medical images. *Computers & Security, 99*, Article 102021. https://doi.org/10.1016/j.cose.2020.102021

Haque, R. U., Hasan, A. S. M. T., Daria, A., Rasool, A., & Chen, H. (2023). A novel secure and distributed architecture for a privacy-preserving healthcare system. *Journal of Network and Computer Applications*, 103696. https://doi.org/10.1016/j.jnca.2023.103696

Haripriya, R., Khare, N., & Pandey, M. (2025). Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings. *Scientific Reports, 15*(1), Article 12482. https://doi.org/10.1038/s41598-025-97565-4

Hasan, M. M., & Rahman, M. M. (2025). Privacy-preserving polyp segmentation using federated learning with differential privacy. *Smart Health, 36*, Article 100551. https://doi.org/10.1016/j.smhl.2025.100551

He, T., Han, P., Duan, S., Wang, Z., Wu, W., Liu, C., & Han, J. (2024). Generative data augmentation with differential privacy for non-IID problem in decentralized clinical machine learning. *Future Generation Computer Systems, 160*, 171–184. https://doi.org/10.1016/j.future.2024.04.028

Hennebelle, A., Ismail, L., Materwala, H., Al Kaabi, J., Ranjan, P., & Janardhanan, R. (2024). Secure and privacy-preserving automated machine learning operations into end-to-end integrated IoT-edge-artificial intelligence-blockchain monitoring system for diabetes mellitus prediction. *Computational and Structural Biotechnology Journal, 23*, 212–233. https://doi.org/10.1016/j.csbj.2023.11.038

Hossen, M. N., Panneerselvam, V., Koundal, D., Ahmed, K., Bui, F. M., & Ibrahim, S. M. (2023). Federated machine learning for detection of skin diseases and enhancement of Internet of Medical Things (IoMT) security. *IEEE Journal of Biomedical and Health Informatics, 27*(2), 835–841. https://doi.org/10.1109/JBHI.2022.3149288

Houser, S. H., Flite, C. A., & Foster, S. L. (2023). Privacy and security risk factors related to telehealth services - A systematic review. *Perspectives in Health Information Management, 20*(1), 1f.

Huang, T., Xu, J., Tu, S., & Han, B. (2023). Robust zero-watermarking scheme based on a depthwise overparameterized VGG network in healthcare information security. *Biomedical Signal Processing and Control, 81*, Article 104478. https://doi.org/10.1016/j.bspc.2022.104478

Ibrahim, M., Al-Wadi, A., & Elhafiz, R. (2024). Security analysis for smart healthcare systems. *Sensors, 24*(11), Article 3375. https://doi.org/10.3390/s24113375

Inam, S., Kanwal, S., Anwar, A., Mirza, N. F., & Alfraihi, H. (2024). Security of End-to-End medical images encryption system using trained deep learning encryption and decryption network. *Egyptian Informatics Journal, 28*, Article 100541. https://doi.org/10.1016/j.eij.2024.100541

Iqbal, S., Choudhry, I. A., Ullah, I., Kaur, K., Choi, B. J., & Hassan, M. M. (2025). Domain adaptive FL for edge-enabled privacy-preserving MRI analysis. *Alexandria Engineering Journal, 128*, 324–339. https://doi.org/10.1016/j.aej.2025.02.048

Jackson, G., Hu, J., & Section Editors for the IMIA Yearbook Section on Artificial Intelligence in Health (2019). Artificial Intelligence in Health in 2018: New Opportunities, Challenges, and Practical Implications. *Yearbook of Medical Informatics*, *28*(1), 52–54. https://doi.org/10.1055/s-0039-1677925

John, S., & Kumar, S. N. (2023a). IoT based medical image encryption using linear feedback shift register – Towards ensuring security for teleradiology applications. Measurement: Sensors, 25, Article 100676. https://doi.org/10.1016/j.measen.2023.100676

John, S., & Kumar, S. N. (2023b). 2D Lorentz chaotic model coupled with logistic chaotic model for medical image encryption: Towards ensuring security for teleradiology. *Procedia Computer Science, 218*, 918–926. https://doi.org/10.1016/j.procs.2023.01.072

Kaissis, G., Ziller, A., Passerat-Palmbach, J., Ryffel, T., Usynin, D., Trask, A., Lima, I., Jr., Mancuso, J., Jungmann, F., Steinborn, M.-M., Saleh, A., Makowski, M., Rueckert, D., & Braren, R. (2021). End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nature Machine Intelligence*, *3*(6), 473–484. https://doi.org/10.1038/s42256-021-00337-8

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine, 17*(1), Article 195. https://doi.org/10.1186/s12916-019-1426-2 (scirp.org)

Khullar, D., Casalino L.P., Qian, Y., Lu, Y., Krumholz, H.M., & Aneja, S. (2022). Perspectives of patients about artificial intelligence in health care. *JAMA Network Open, 5(*5): e2210309. doi:10.1001/jamanetworkopen.2022.10309

Kokin, H., Lytvyn, O., & Nguyen, G. (2024). Optimal differential privacy for deep learning model training. *Procedia Computer Science*, *246*, 2419–2428. https://doi.org/10.1016/j.procs.2024.09.483

Kolhar, M., & Misfer, M. (2023). An intelligent cardiovascular diseases prediction system focused on privacy. *Intelligent Automation & Soft Computing*, *36*(1), 529–542. https://doi.org/10.32604/iasc.2023.030098

Lee, T.-H., Kim, S., Lee, J., & Jun, C.-H. (2024). HarmoSATE: Harmonized embedding-based self-attentive encoder to improve accuracy of privacy-preserving federated predictive analysis. *Information Sciences*, *662*, Article 120265. https://doi.org/10.1016/j.ins.2024.120265

Li, L., Peng, W., & Rheu, M. M. J. (2024). Factors predicting intentions of adoption and continued use of artificial intelligence chatbots for mental health: Examining the role of UTAUT model, stigma, privacy concerns, and artificial intelligence hesitancy. *Telemedicine and e-Health*, *30*(3), https://doi.org/10.1089/tmj.2023.0313

Li, H., Zhao, X., Liu, H., Wang, Z., Cai, Y., Yang, C., & Cong, F. (2025). Distributed data-privacy preserving federated learning method for sleep stage classification. *Biomedical Signal Processing and Control*, *109*, Article 108032. https://doi.org/10.1016/j.bspc.2025.108032

Li, Y.-H., Li, Y.-L., Wei, M.-Y., & Li, G.-Y. (2024). Innovation and challenges of artificial intelligence technology in personalized healthcare. *Scientific Reports, 14*, Article 18994. https://doi.org/10.1038/s41598-024-70073-7

Liu, K., & Tao, D. (2022). The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. *Computers in Human Behavior*, *127*, Article 107026. https://doi.org/10.1016/j.chb.2021.107026

Montenegro, H., Silva, W., & Cardoso, J. S. (2021). Privacy-preserving generative adversarial network for case-based explainability in medical image analysis. *IEEE Access*, *9*, 148037-148047. https://doi.org/10.1109/ACCESS.2021.3124844

Moy, S., Irannejad, M., Jeanneret Manning, S., Farahani, M., Ahmed, Y., Gao, E., Prabhune, R., Lorenz, S., Mirza, R., & Klinger, C. (2024). Patient perspectives on the use of artificial intelligence in health care: A scoping review. *Journal of Patient-Centered Research and Reviews, 11*(1), 51–62. https://doi.org/10.17294/2330-0698.2029

National Academy of Medicine. (2025). *An artificial intelligence code of conduct for health and medicine: Essential guidance for aligned action*. Washington, DC: National Academies Press. https://doi.org/10.17226/29087

Nowak, S., Wulff, B., Layer, Y. C., Theis, M., Isaak, A., Salam, B., Block, W., Kuetting, D., Pieper, C.C., Luetkens, J.A., Attenberger, U.& Sprinkart, A. M. (2025). Privacy-ensuring open-weights large language models are competitive with closed-weights GPT-4o in extracting chest radiography findings from free-text reports. *Radiology*, *314*(1), e240895. https://doi.org/10.1148/radiol.240895

Om Kumar, C.U, Gajendran, S., Bhavadharini, R. M., Suguna, M. (2023). EHR privacy preservation using federated learning with DQRE-Scnet for healthcare application domains. *Knowledge-Based Systems*, *275*. https://doi.org/10.1016/j.knosys.2023.110638

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Aki, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine, 18*(3), e1003583. https://doi.org/10.1371/journal.pmed.1003583

Ragab, M., Bahaddad, A.A., Hamed, D., et al. (2024). Blockchain-driven privacy preserving electronic health records analysis using sine cosine algorithm with deep learning model. *Human-centric Computing and Information Sciences, 14*(9). https://doi.org/10.22967/HCIS.2024.14.009

Sangaiah, A. K., Javadpour, A., Ja'fari, F., Pinto, P., & Chuang, H. M. (2023). Privacy-aware and AI techniques for healthcare based on k-anonymity model in internet of things. *IEEE Transactions on Engineering Management*, *71*, 12448-12462.

Seyfizadeh, A., Peach, R. L., Tovote, P., Isaias, I. U., Volkmann, J., & Muthuraman, M. (2024). Enhancing security in brain–computer interface applications with deep learning: Electroencephalogram-based user identification. *Expert Systems with Applications*, *253*, 124218. https://doi.org/10.1016/j.eswa.2024.124218

Shao, R., He, H., Chen, Z., Liu, H., & Liu, D. (2020). Stochastic channel-based federated learning with neural network pruning for medical data privacy preservation: Model development and experimental validation. *JMIR Formative Research*, *4*(12), e17265. https://doi.org/10.2196/17265

Shojaei, P., Vlahu-Gjorgievska, E., & Chow, Y.-W. (2024). Security and privacy of technologies in health information systems: A systematic literature review. *Computers, 13*(2), 41. https://doi.org/10.3390/computers13020041

Singh, N., Buyya, R., & Kim, H. (2025). Securing cloud-based internet of things: Challenges and mitigations. *Sensors, 25(*1):79. https://doi.org/10.3390/s25010079

Son, Y., Han, K., Lee, Y. S., Yu, J., Im, Y. H., & Shin, S. Y. (2021). Privacy-preserving breast cancer recurrence prediction based on homomorphic encryption and secure two-party computation. PLOS One, 16(12), e0260681. https://doi.org/10.1371/journal.pone.0260681

Supraja, B., Kumar, V. K., & Kumar, N. K. (2025). DARL: Effectual Deep Adaptive Reinforcement Learning Model enabled security and energy-efficient healthcare system in IoT with Modified Manta Ray Foraging Optimization. *Biomedical Engineering: Applications, Basis and Communications, 37*(6), 2550012. https://doi.org/10.4015/S1016237225500127

Tang, S., Cheang, C. F., Yu, X., Liang, Y., Feng, Q., & Chen, Z. (2023). TransCS-Net: A hybrid transformer-based privacy-protecting network using compressed sensing for medical image segmentation. *Biomedical Signal Processing and Control*, *86*, 105131. https://doi.org/10.1016/j.bspc.2023.105131

Tang, Z., Wong, H. S., & Yu, Z. (2024). Privacy-preserving federated learning with domain adaptation for multi-disease ocular disease recognition. *IEEE Journal of Biomedical and Health Informatics*, *28*(6), 3219-3227.

The Joint Commission. (2025). *The responsible use of AI in healthcare* Guidance document]. The Joint Commission. https://digitalassets.jointcommission.org/api/public/content/dcfcf4f1a0cc45cdb526b3cb034c68c2

Teo, Z. L., Zhang, X., Yang, Y., Jin, L., Zhang, C., Poh, S. S. J., Yu, W., Chen, Y., Jonas, J. B., Wang, Y. X., Wu, W.-C., Lai, C.-C., Liu. Y., Goh, R. S. M, & Ting, D. S. W. (2025). Privacy-preserving technology using federated learning and blockchain in protecting against adversarial attacks for retinal imaging. *Ophthalmology*, *132*(4), 484-494. https://doi.org/10.1016/j.ophtha.2024.10.017

Tsoi, A. H., Gartner, G., Cotten, S. W., Kim, J., Nazarian, J., Thomas, J., McSwain, S. D., Ahmadi-Moosavi, R., & Rimal, R. (2025). Establishing and implementing a responsible artificial intelligence framework: A 1-year review. *Journal of the American Medical Informatics Association, 32*(11), 1778–1784. https://doi.org/10.1093/jamia/ocaf147

U.S. Department of Health & Human Services. (2023a). *HIPAA privacy rule*. https://www.hhs.gov/hipaa/for-professionals/privacy/index.html

U.S. Department of Health & Human Services. (2023b). *HIPAA security rule*. https://www.hhs.gov/hipaa/for-professionals/security/index.html

Veritas Health Innovation. (n.d.). *Covidence systematic review software* (Version 2). https://www.covidence.org

Vignesh, S., Mathi, S., Dayanand, V., & Ramesh, G. (2025). Privacy-aware reconstruction and profiling using advanced contextual embeddings. *Procedia Computer Science*, *258*, 578-587. https://doi.org/10.1016/j.procs.2025.04.292

Vizitiu, A., Nită, C. I., Puiu, A., Suciu, C., & Itu, L. M. (2019, July). Privacy-preserving artificial intelligence: Application to precision medicine. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 6498-6504). IEEE.

Wang, J., Quasim, M. T., & Yi, B. (2025). Privacy-preserving heterogeneous multi-modal sensor data fusion via federated learning for smart healthcare. *Information Fusion*, *120*, 103084. https://doi.org/10.1016/j.inffus.2025.103084

Wang, Q. (2025). Interpretable vertical federated learning with privacy-preserving multi-source data integration for prognostic prediction. *Engineering Applications of Artificial Intelligence*, *148*, 110408. https://doi.org/10.1016/j.engappai.2025.110408

Wang, S., Yang, F., Wang, S., & Sun, R. (2025). Integrating differential privacy in deep reinforcement learning for sepsis treatment with pulmonary implications. *Frontiers in Medicine*, *12*, 1590824. https://doi.org/10.3389/fmed.2025.1590824

Xu, X., Peng, H., Bhuiyan, M. Z. A., Hao, Z., Liu, L., Sun, L., & He, L. (2021). Privacy-preserving federated depression detection from multisource mobile health data. *IEEE Transactions on Industrial Informatics*, *18*(7), 4788-4797.

Yahiaoui, M. E., Derdour, M., Abdulghafor, R., Turaev, S., Gasmi, M., Bennour, A., Aborujilah, A. & Sarem, M. A. (2024). Federated learning with privacy preserving for multi-institutional three-dimensional brain tumor segmentation. *Diagnostics*, *14*(24), 2891. https://doi.org/10.3390/diagnostics14242891

Yan, C., Zeng, X., Xi, R., Ahmed, A., Hou, M., & Tunio, M. H. (2023). PLA—A Privacy-embedded lightweight and efficient automated breast cancer accurate diagnosis framework for the internet of medical things. *Electronics*, *12*(24), 4923. https://doi.org/10.3390/electronics12244923

Yang, J., Gu, Y., Gao, S., Ren, W., & Yu, Z. (2025). Fed-GFM-DG: A privacy-preserving framework for fundus image segmentation via generative-based feature generalization and mask-guided aggregation. *Biomedical Signal Processing and Control*, *106*, 107725. https://doi.org/10.1016/j.bspc.2025.107725

Zarifis, A., Kawalek, P., & Azadegan, A. (2021). Evaluating if trust and personal information privacy concerns are barriers to using health insurance that explicitly utilizes AI. *Journal of Internet Commerce*, *20*(1), 66-83. https://doi.org/10.1080/15332861.2020.1832817

Zerka, F., Urovi, V., Vaidyanathan, A., Barakat, S., Leijenaar, R. T., Walsh, S., Gabrani-Juma, H., Miraglio, B., Woodruff, H. C., Dumontier, M. & Lambin, P. (2020). Blockchain for privacy preserving and trustworthy distributed machine learning in multicentric medical imaging (C-DistriM). *IEEE Access*, *8*, 183939-183951. https://doi.org/10.1109/ACCESS.2020.3029445

Zhang, L., Fang, G., & Tan, Z. (2025). FedCCW: A privacy-preserving Byzantine-robust federated learning with local differential privacy for healthcare. *Cluster Computing*, *28*(3), 182. https://doi.org/10.1007/s10586-024-04894-6

Zhang, Y., Kohne, J. G., Webster, K., Vartanian, R., Wittrup, E., & Najarian, K. (2025). AXpert: Hduman expert facilitated privacy-preserving large language models for abdominal X-ray report labeling. *JAMIA Open*, *8*(1). https://doi.org/10.1093/jamiaopen/ooaf008

Zhang, X., Wang, Y., Ma, J., & Jin, Q. (2022). QAPP: A quality-aware and privacy-preserving medical image release scheme. *Information Fusion*, *88*, 281-295. https://doi.org/10.1016/j.inffus.2022.07.011

HAWAI'I PACIFIC UNIVERSITY LIBRARY & LEARNING COMMONS

This work is published by Hawaii Pacific University Library & Learning Commons and is licensed under a Creative Commons Attribution 4.0 International License.

Appendix

Table A1

Summary Table of Research Journals

(Authors, Year) *Journal*	Privacy	Security	Country/ Region	Study Design	Study Site
(Aanjankumar et al., 2025) *IEEE Access*	Privacy-preserving healthcare data processing framework using differential privacy		India	Experimental	Hospital
(Abdalla et al., 2020) *Journal of Medical Internet Research*	Evaluated privacy risks of word embeddings trained on deidentified clinical notes, and found that sensitive patient information can be reconstructed from embeddings even after PHI removal.		Canada	Experimental	Clinic; Hospital
(Ain et al., 2023) *Diagnostics*	Proposed a federated learning–based skin cancer prediction framework that improves classification accuracy while preserving patient data privacy by avoiding centralized data collection.		Saudi Arabia	Case Study	Other
(Akmese, 2024) *BMC Medical Informatics and Decision Making*		Utilized encrypted biomarkers to ensure confidentiality to predict individuals with pancreatic ductal adenocarcinoma	Turkey	Case-Control	Hospital
(Alabdulwahab et al., 2024) *Sensors*	Synthetic data generation methods using a conditional tabular generative adversarial network (CT-GAN) aimed at maintaining the utility of IoT sensor network data for intrusion-detection systems (IDS) while safeguarding privacy.		Korea	Experimental	Non-healthcare Setting
(Almaiah et al., 2023) *Electronics*		Used a combination of IoTs data (EHRs, smart phones, smart watches, etc.) and an Encrypted Dynamic Bayesian Network)	Jordan, India & Saudi Arabia	Pilot Study	Other

		technique to secure the communication protocols so that only authorized people could access the data			
(Almufareh, 2024) *IEEE Access*	Ensures privacy by performing real-time analysis and classification of skin lesion images directly on edge devices, minimizing data transmission and safeguarding sensitive patient information.		Saudi Arabia	Experimental	Clinic; Hospital
(Amoon et al., 2020) *Measurement*	Addressing privacy and security risks in healthcare data transmission and storage.	Applying AI-based spiral swarm optimized artificial immune system (SOAIS) to detect threats and ensure data integrity. Evaluating performance against malware and unauthorized access.	Saudi Arabia	Experimental	Other
(Anita et al., 2025) *Computer Methods and Programs in Biomedicine*		Developed multimodal user authentication.	India	Model Development	Non-healthcare Setting
(Arafat et al., 2025) *IET Blockchain*	Presents a blockchain-based framework for secure, scalable, and efficient sharing of medical images, prioritizing data integrity and privacy.	Presents a blockchain-based framework for secure and efficient sharing of medical images, prioritizing data integrity and privacy.	Bangladesh & Finland	Experimental	Non-healthcare Setting
(Arasteh et al., 2024) *Radiology: Artificial Intelligence*		Studied privacy-safe medical AI performance for X-rays	Germany	Experimental	Clinic; Hospital
(Bai et al., 2025) *Computer Methods and Programs in Biomedicine*		Propose PPCNN, a practical and privacy-preserving framework for CNN Inference, to enable secure sharing of body-related radiological images for deep learning.	China	Model Development	Non-healthcare Setting
(Bangare et al., 2024) *Frontiers in Health Informatics*	Evaluated federated transfer learning models under IID and non-IID conditions, highlighting federated learning's effectiveness for distributed medical data without centralizing sensitive information.		India	Experimental	Non-healthcare Setting

(Barnawi et al., 2024) *IEEE Transactions on Network and Service Management*	Proposed a framework that combines Federated Learning (FL) and Differential Privacy (DP) to preserve the privacy of chest X-ray datasets used to develop and analyze high-performing Convolutional Neural Networks (CNNs) for detecting Tuberculosis.		Saudi Arabia	Experimental	Non-healthcare Setting
(Baumgartner et al., 2023) *Journal of Healthcare Informatics Research*	Compared decentralized learning methods for ECG classification and demonstrated that federated learning provides strong privacy protection with minimal performance loss.		Austria	Pilot Study	Non-healthcare Setting
(Bavli et al., 2025) *AI & Society*	Ethical implications of AI health monitoring for Parkinson's disease. Stakeholder views on data privacy, ownership, and access to health data, use and potential misuse of data, and personal privacy at home.	Ethical implications of AI health monitoring for Parkinson's disease. Stakeholder views on security, ownership, and access to health data, use and potential misuse of data, and personal privacy at home.	Canada	Qualitative	Other
(Benouis et al., 2024) *JMIR Mental Health*	To address privacy concerns with the sensitive data collected by wearable sensors, a framework that integrates differential privacy and federated learning approaches with Multi-Task Learning (MTL was used.		Germany	Experimental	Non-healthcare Setting
(Bergen et al., 2024) *Computer Methods and Programs in Biomedicine*	Proposed a 3-D generative model, Transversal GAN (TrGAN), to generate medical images that retain statistical properties and image features of the training data while balancing privacy and utility.		Canada	Experimental	Other
(Bharath et al., 2024) *The Open Public Health Journal*		Employed wavelet transform techniques to compress the Region of Interest (ROI) within medical images to retain the integrity of diagnostic information while reducing file size, improving performance for telemedicine.	India	Experimental	Clinic; Hospital

(Bi et al., 2022) *IEEE Transactions on Industrial Informatics*		Separate private data, analyze health information securely in a cloud system, and construct a security module based on the convolutional neural network.	China	Pilot Study	Other
(Butt et al., 2023) *Electronics*	Proposed a collaborative federated learning architecture for COVID-19 screening via chest X-ray images that preserves patient privacy by keeping imaging data localized across institutions.		Saudi Arabia	Model Development	Non-healthcare Setting
(Om Kumar C.U. et al., 2023) *Knowledge-Based Systems*		Designed a federated learning (FL) framework for secure EHR sharing across hospitals	India	Experimental	Other
(Cao et al., 2023) *Computers & Security*	Proposed a privacy-preserving healthcare monitoring framework and interactive protocols based on long short-term memory (LSTM) for edge servers.		China	Experimental	Non-healthcare Setting
(Chen, Cheng, et al., 2024) *IEEE Access*		Proposed a symmetric encryption and decryption protocol to ensure the security of biosignals and medical images and assist in disease diagnosis using a small-scale cascaded CNN architecture.	Taiwan	Experimental	Hospital
(Chen, Song, et al., 2024) *Image and Vision Computing*	Introduced Privacy-SF, an encoding-based framework for privacy-preserving medical image segmentation that protects sensitive features during both training and inference.		China	Experimental	Hospital
(Choudhury et al., 2025) *JMIR AI*	Developed and demonstrated a federated deep learning infrastructure—the Personal Health Train (PHT)—for training deep learning models on medical imaging data without sharing patient-level data across institutions.		Netherlands	Infrastructure Evaluation	Clinic; Hospital
(Chuang et al., 2025) *Journal of the American*	Evaluated privacy risks and data utility trade-offs in using LLMs—GPT-3.5, GPT-4,		United States & Canada	Experimental	Hospital

Medical Informatics Association	and Mistral 7B—to generate synthetic EHR notes.				
(Deepak et al., 2024) *Measurement: Sensors*	Federated deep learning (FL) framework detected stress from wearable heart activity signals with accuracy comparable to traditional centralized methods (FL: 81.75%, centralized: 81.65%) while preserving privacy.		India	Experimental	Non-healthcare Setting
(Dubey et al., 2022) *International Journal of Electronics and Telecommunications*		Sensitive data stays local; advanced privacy methods ensure strong compliance with regulations.	India	Experimental	Other
(Dubey et al., 2025) *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*	Used federated learning for real-time mental health diagnostics by integrating multimodal data (physiological signals, online activity, social media interactions).		India & South Africa	Experimental	Non-healthcare Setting
(Ellis et al., 2025) *Journal of the American Geriatrics Society*	Privacy concerns regarding information shared with novel patient-facing chatbots.		United States	Mixed Methods	Clinic; Hospital
(Ferdowski et al., 2024) *Computer Methods and Programs in Biomedicine*		Uses differential privacy and data anonymization to protect sensitive attributes, ensuring secure and ethical AI-based cardiovascular disease prediction.	United States, Hungary, & Switzerland	Experimental	Other
(Gong et al., 2025) *IEEE Journal of Biomedical and Health Informatics*	Experimental performance evaluation using real-world cases from medical question-answering datasets; comparison with existing large language models for accuracy.		China	Case Study	Non-healthcare Setting
(Guan et al., 2023) *International Journal of Surgery*	Presented machine learning (ML) models to predict the lymph node metastasis (LNM) status for rectal cancer patients before surgery based on a privacy-preserving computing platform (PPCP).		China & United States	Cohort	Hospital
(Gulati et al., 2025) *International Journal of*	Applied federated learning to ocular disease detection, demonstrating accurate diagnosis		China	Experimental	Non-healthcare Setting

Mathematical, Engineering and Management Services	while maintaining patient privacy by avoiding centralized data storage.				
(Guo et al., 2020) *Computers & Security*	Proposed a classification and retrieval method to search for and locate relevant cases by constructing a privacy-preserving Convolutional Neural Network (CNN) framework that allows the classification and searching of secure, content-based, large-scale encrypted images with homomorphic encryption.		China	Experimental	Non-healthcare Setting
(Haque et al., 2023) *Journal of Network and Computer Applications*	Security model development and testing.		China	Model Development	Non-healthcare Setting
(Haripriya et al., 2025) *Scientific Reports*	Integrated transfer learning with federated learning to enable scalable and privacy-preserving medical image classification across multiple institutions.		India	Model Development	Hospital; Other
(Hasan et al., 2025) *Smart Health*	Designed a differentially private federated learning system for polyp segmentation that achieves high accuracy without exposing sensitive patient data.		Bangladesh	Model Development	Hospital
(He et al., 2024) *Future Generation Computer Systems*	Proposed a differentially private swarm learning framework using generative augmentation to address non-IID clinical data while providing formal privacy guarantees.		China	Experimental	Clinic; Hospital
(Hennebelle et al., 2024) *Computational and Structural Biotechnology Journal*		Developed a secure, privacy-preserving, end-to-end system integrating IoT, edge computing, AI, and blockchain for predicting type 2 diabetes based on risk factors.	Australia, United Arab Emirates, & India	Experimental	Non-healthcare Setting
(Hossen et al., 2023) *IEEE Journal of Biomedical and Health Informatics*	Evaluated a federated learning–based CNN framework for skin disease classification that ensures patient data privacy while improving model generalization across distributed clients.		New Zealand	Pilot Study	Clinic; Hospital

(Huang et al., 2023) *Biomedical Signal Processing and Control*	Proposed a deep learning–based zero-watermarking scheme for medical images that enhances protection against tampering and attacks through encryption and robust watermark security.		China	Experimental	Clinic; Hospital
(Ibrahim et al., 2024) *Sensors*	Implementation and evaluation of an ensemble intrusion detection system using honeypot diversion and machine learning.		Jordan	Experimental	Non-healthcare Setting
(Inam et al., 2024) *Egyptian Informatics Journal*	Proposed Cycle_GAN (Cycle-Consistent Generative Adversarial Network) to protect the confidentiality of unpaired medical images used in deep learning.		Pakistan & Saudi Arabia	Experimental	Non-healthcare Setting
(Iqbal et al., 2025) *Alexandria Engineering Journal*	Introduced a domain-adaptive federated learning framework combined with edge computing to enable privacy-preserving multi-modal medical image analysis across heterogeneous institutions.		China, Pakistan, Korea, Canada, India, & Saudi Arabia	Experimental	Non-healthcare Setting
(John & Kumar, 2023a) *Measurement: Sensors*	Proposed an encryption scheme using a linear feedback shift register (LFSR) for DICOM CT images to be securely and efficiently shared through the cloud for telemedicine applications.		Malaysia & India	Algorithmic Development	Non-healthcare Setting
(John et al., 2023b) *Procedia Computer Science*	Proposes a hybrid chaotic model combining the 2D Lorentz Chaotic model with the Logistic chaotic model for secure encryption/decryption of DICOM CT images through the cloud to improve teleradiology applications.		Malaysia	Experimental	Other
(Kaissis et al., 2021) *Nature Machine Intelligence*	Presented PriMIA (Privacy-preserving Medical Image Analysis), a framework for differentially private, securely aggregated federated learning and encrypted inference on medical imaging data.		Germany	Case Study	Other
(Kokin et al., 2024) *Procedia Computer Science*	Investigated how differential privacy can optimize both privacy and model performance during deep-learning training,		Slovakia	Model Development	Non-healthcare Setting

	specifically in sensitive medical data applications (heart disease).				
(Kolhar & Misfer, 2023) *Intelligent Automation & Soft Computer*	Proposes a process to combine machine learning and cloud computing to protect patient and provider data when sharing heart rate data streams to predict myocardial infarction.		Saudi Arabia	Experimental	Non-healthcare Setting
(Lee et al., 2024) *Information Sciences*	Federated learning framework (HarmoSATE) that improved predictive accuracy using harmonized contextual embeddings and self-attention mechanisms across EHR data		Republic of Korea	Experimental	Hospital
(Li, Peng, et al., 2024) *Telemedicine and e-Health*	Investigates predictors of adoption and continued use of AI chatbots for mental health among U.S. adults with depression and anxiety symptoms.		United States	Cross-sectional	Other
(Li et al., 2025) *Biomedical Signal Processing and Control*	Demonstrated that federated learning enables accurate sleep stage classification while protecting patient data and mitigating class imbalance across distributed clients.		China & Finland	Experimental	Clinic; Hospital
(Liu et al., 2022) *Computers in Human Behavior*	Investigates how trust and AI-related features affect behavioral intention to use smart healthcare services, and how these relationships vary by gender, age, and usage experience.		China	Cross-sectional	Other
(Montenegro et al., 2021) *IEEE Access*	Proposed a privacy-preserving generative adversarial network for the privatization of case-based explanations, guaranteeing privacy, intelligibility, and preservation of explanatory evidence using interpretability saliency maps to reconstruct relevant features.		Portugal	Model Development	Non-healthcare Setting
(Nowak et al., 2025) *Radiology*	Publicly available and privacy ensuring LLMs can perform just as well as expensive proprietary systems like GPT-4o for		Germany	Exploratory	Hospital

	organizing information from doctors' radiology reports.				
(Ragab et al., 2024) *Human-centric Computing and Information Sciences*	Development of a blockchain privacy preserving technique for EHR analysis and for storing of patients' clinical and institutional data.		Saudi Arabia	Model Development	Non-healthcare Setting
(Sangaiah et al., 2024) *IEEE Transactions on Engineering Management*	Enhanced privacy protection in healthcare data mining by applying k-anonymity and entropy-based strategies to anonymize sensitive patient information in distributed IoT environments.		Taiwan	Algorithmic Development	Non-healthcare Setting
(Seyfizadeh et al., 2024) *Expert Systems with Applications*		Used a deep residual neural network (ResNet) for the identification of individuals based on their EEG signals, to enhance security in brain-computer interface (BCI) applications through improved user identification accuracy and superior performance over other methods.	Germany	Experimental	Non-healthcare Setting
(Shao et al., 2020) *JMIR Formative Research*	Developed and validated a privacy-preserving federated learning method that enables distributed model training without sharing sensitive medical data.		China & U.S.	Experimental	Other
(Son et al., 2021) *PLoS ONE*	Proposed a method for privacy-preserving Gated Recurrent Unit (GRU) inference model using privacy-enhancing technologies, including homomorphic encryption and secure two-party computation in the prediction of breast cancer.		South Korea	Cohort	Other
(Supraja et al., 2025) *Biomedical Engineering Applications Basis and Communications*	.	Secure transmission of patient data via IoT sensors	India	Experimental	Non-healthcare Setting

(Tang et al., 2023) *Biomedical Signal Processing and Control*	Proposed a compressed sensing–based transformer network for medical image segmentation that reduces data exposure while maintaining performance under strict data protection regulations.		China	Model Development	Non-healthcare Setting
(Tang et al., 2024) *IEEE Journal of Biomedical and Health Informatics*	Developed a privacy-preserving federated learning framework with domain adaptation for ocular disease recognition using Gaussian noise mechanisms to protect local data.		Hong Kong & China	Model Development	Non-healthcare Setting
(Teo et al., 2025) *Ophthalmology*	Proposed federated learning algorithm applying blockchain to overcome privacy shortcomings when transferring images for the detection of myopic macular degeneration.		China	Cohort	Non-healthcare Setting
(Vignesh et al., 2025) *Procedia Computer Science*	AI models (including BERT, a language processing system) combined with autoencoders to achieve 93% accuracy in identifying mental health conditions while keeping PHI protected.		India	Experimental	Other
(Vizitiu et al., 2019) *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*	Proposed a solution, based on the MORE homomorphic encryption scheme, to allow deep learning models to maintain confidentiality of medical image data while enabling computations.		Romania	Model Development	Non-healthcare Setting
(Wang, S. et al., 2025) *Frontiers in Medicine*	Developed a differential privacy–based deep reinforcement learning (DP-DRL) framework for sepsis treatment, particularly in patients with pulmonary complications		China	Experimental	Non-healthcare Setting
(Wang, Q., 2025) *Engineering Applications of Artificial Intelligence*	Proposed an interpretable vertical federated learning framework that integrates private multi-source data to improve cancer prognosis prediction accuracy without sharing raw data.		China	Experimental	Non-healthcare Setting

(Wang, J., et al., 2025) *Information Fusion*	Introduced PHMS-Fed, a privacy-preserving federated learning framework for heterogeneous multi-modal sensor fusion in smart healthcare systems.		China	Experimental	Non-healthcare Setting
(Xu et al., 2022) *IEEE Transactions on Industrial Informatics*	Developed a federated learning framework for detecting depression using multisource mobile health data while preserving patient privacy.		China	Experimental	Non-healthcare Setting
(Yahiaoui et al., 2024) *Diagnostics*	Presented a privacy-aware federated learning framework with differential privacy techniques for distributed brain tumor segmentation without compromising data confidentiality.		Algeria, Oman, United Arab Emirates, & Yemen	Experimental	Non-healthcare Setting
(Yan et al., 2023) *Electronics*	Proposed a lightweight privacy-embedded federated learning model for breast tumor detection in IoMT environments that improves diagnostic accuracy while protecting patient data.		China	Experimental	Non-healthcare Setting
(Yang et al., 2025) *Biomedical Signal Processing and Control*	Developed a privacy-preserving federated domain generalization framework for multi-center eye imaging using synthetic data generation and secure global aggregation to improve model robustness across domains.		China	Experimental	Non-healthcare Setting
(Zarifis et al., 2021) *Journal of Internet Commerce*	Investigates how visible AI involvement in health insurance purchasing affects: trust in the system and privacy concerns about sharing sensitive health and financial data.		United Kingdom	Cross-sectional	Other
(Zerka et al., 2020) *IEEE Access*	Demonstrated effectiveness of Chained Distributed Machine learning (C-DistriM), combining sequential distributed learning with blockchain to provide transparency and increase trust in AI computations on medical images.		Netherlands & Belgium	Case Study	Hospital
(Zhang et al., 2022) *Information Fusion*	Proposes an image release scheme, QAPP, which integrates the discrete cosine		China	Experimental	Hospital

	transform (DCT) with differential privacy (DP) to maintain confidentiality of high-quality medical images.				
(Zhang et al., 2025) *JAMIA Open*	Developed and evaluated LLMs (Axpert) for automated labeling of pediatric abdominal X-ray (AXR) reports, for necrotizing enterocolitis (NEC) diagnosis, and subtype classification shows that Axpert outperforms baseline models.		United States	Qualitative	Hospital
(Zhang et al., 2025) *Cluster Computing*	Introduced FedCCW, a Byzantine-robust and privacy-preserving federated learning scheme using differential privacy and secure aggregation to enable safe multi-institutional medical collaboration.		China	Experimental	Non-healthcare Setting