Review article

DIGITAL HEALTH

DIGITAL HEALTH
Volume 11: 1–19

DOI: 10.1177/20552076251343959
journals.sagepub.com/home/dhj

Sage

Data privacy in healthcare: Global challenges and solutions

Andrew Kweku Conduah[1,2], Sebastian Ofoe[3] and Dorothy Siaw-Marfo[4]

Abstract

Purpose: This study explores global frameworks for healthcare data privacy, focusing on the General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), and the Protection of Personal Information Act (POPIA). It examines the challenges of regional regulatory disparities, systemic vulnerabilities identified through major health data breach case studies, and the potential of advanced technologies to enhance privacy protections.

Methods: A qualitative research approach was adopted, incorporating corpus construction and comparative analysis of legal and technical frameworks. The study also utilized case studies of significant health data breaches to identify vulnerabilities and evaluate the role of emerging technologies, such as artificial intelligence (AI) and machine learning (ML), in mitigating risks and enhancing regulatory compliance.

Results: Findings indicate that GDPR, CCPA, and POPIA set high standards for data protection but reveal significant variability in enforcement and technological adoption across regions. Challenges include inconsistent definitions of sensitive data, semantic discrepancies, a lack of standardized protocols, and limited information technology infrastructure in certain jurisdictions. Advanced technologies like AI and ML promise to address these gaps by improving data harmonization and security.

Conclusions: Addressing healthcare data privacy challenges requires harmonized global regulations, advanced technological tools, and international collaboration. Strengthening frameworks, enhancing information technology infrastructure, and employing semantic models and ontologies are essential for protecting sensitive data, ensuring compliance, and fostering public trust in digital healthcare systems.

Keywords

Data privacy, healthcare, General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA), Protection of Personal Information Act (POPIA), data security

Received: 4 February 2025; accepted: 5 May 2025

Introduction

Data privacy in healthcare entails protecting sensitive patient information, including medical records, personal identifiers, and other health-related data, from unauthorized access, misuse, or disclosure.[1,2] To clarify, "personal data" refers to any information that can identify an individual (such as names, addresses, and contact details), whereas "medical data" specifically pertains to information generated during healthcare delivery, such as diagnoses, treatment records, and clinical outcomes. Although these categories often intersect, medical data is intimately linked to patient care and is subject to additional ethical and professional safeguards imposed by the deontological codes of healthcare providers.[3] Accordingly, the term "personal

[1]Department of Business Administration, Institute of Work, Employment & Society (IWES), University of Professional Studies (UPSA), Accra, Ghana
[2]Regional Institute for Population Studies (RIPS), University of Ghana, Accra, Ghana
[3]Joshua Alabi Library, Department of Electronic Information Resources, University of Professional Studies, Accra (UPSA), Accra, Ghana
[4]Institute of Work, Employment & Society (IWES), University of Professional Studies (UPSA), Accra, Ghana

Corresponding author:
Andrew Kweku Conduah, Department of Business Administration, Institute of Work, Employment & Society (IWES), University of Professional Studies (UPSA), Post Office Box LG 149, Legon, Accra, Ghana; Regional Institute for Population Studies (RIPS), University of Ghana, P.O. Box LG 96, Legon, Accra, Ghana.
Emails: andrew.conduah@upsamail.edu.gh; akconduah001@st.ug.edu

health information (PHI)" is used in this review to denote data directly related to an individual's health status and care, warranting enhanced protection under both legal regulations and ethical standards.

The World Health Organization (WHO) defines healthcare data privacy as the implementation of measures that guarantee the confidentiality, integrity, and availability of patient information.[4] Similarly, the United Nations Educational, Scientific, and Cultural Organization (UNESCO) describes personal data as any information that directly or indirectly identifies an individual, emphasizing the secure collection, usage, storage, and transfer of such data.[5] In making this distinction, our review aims to underscore that while both personal and medical data require robust safeguards, medical data, by virtue of its connection to patient care, necessitates even stricter protection.

In North America, the US Department of Health and Human Services governs healthcare data privacy through the Health Insurance Portability and Accountability Act (HIPAA), which mandates measures such as access restrictions, encryption protocols, and breach notification requirements for electronically protected health information (ePHI).[6,7] Similarly, Europe's General Data Protection Regulation (GDPR) emphasizes explicit consent, data minimization, and comprehensive technical and organizational safeguards, setting a global benchmark for data protection.[8–10] In the Asia-Pacific region, frameworks like the Asia-Pacific Economic Cooperation (APEC) Privacy Framework advocate for interoperable and culturally tailored privacy protection mechanisms.[11–13] Meanwhile, in sub-Saharan Africa, emerging policies endorsed by the Africa Centers for Disease Control and Prevention (Africa CDC) stress respecting cultural values, securing patient consent, and ensuring accountability in health data management.[14–16] This review examines these diverse regulatory approaches and evaluates how they interact with information systems such as electronic health records (EHRs) and cybersecurity infrastructures to both protect and sometimes compromise PHI.

Despite the presence of robust regulatory frameworks, the increasing integration of EHRs and digital tools has significantly amplified the risk of data breaches and unauthorized access. Notable examples include the Anthem Inc. breach in the USA, which exposed the ePHI of 79 million individuals, and the WannaCry ransomware attack on the UK's National Health Service (NHS), which disrupted critical healthcare services.[17–21] Similarly, the SingHealth breach in Singapore compromised the personal data of 1.5 million patients,[22,23] while incidents in sub-Saharan Africa such as the compromise of the Ghana Health Service's COVID-19 test results portal and the South African medical data breach of 2020 highlight challenges related to inadequate encryption and resource constraints.[24,25] These cases underscore that even with established regulations, systemic vulnerabilities and technological shortcomings can undermine data privacy.

Emerging technologies, particularly blockchain and artificial intelligence (AI), offer promising avenues for addressing these vulnerabilities. Blockchain technology, which operates as a decentralized and immutable digital ledger, can enhance data integrity and transparency by securely recording transactions and preventing unauthorized alterations. Similarly, AI and machine learning (ML) technologies enable real-time breach detection, predictive risk assessment, and automated compliance monitoring. Recent studies, such as those highlighted in Reflections on Blockchain in Health Data Sharing (10.3390/ijerph21020230), demonstrate the transformative potential of these innovations in strengthening healthcare data privacy practices. Balancing technological innovation with stringent privacy protections is, therefore, critical for safeguarding sensitive patient information in an increasingly digitalized healthcare landscape.

This review comprehensively analyzes global healthcare data privacy by examining legal, ethical, and technical dimensions across diverse regulatory frameworks. It emphasizes the need for harmonized global regulations that are adaptable to regional nuances and highlights innovative technological solutions that can bridge current security gaps. Through a detailed analysis of case studies and policies, the study aims to provide actionable insights into enhancing the protection of PHI while maintaining the integrity and trust essential to healthcare delivery.

The study aims to delineate the distinctions between personal and medical data, emphasizing that while both demand robust protection, medical data, being intrinsically tied to patient care, requires additional ethical and professional safeguards. It seeks to examine global regulatory frameworks by analyzing the legal landscapes governing healthcare data privacy in regions such as North America, Europe, Asia-Pacific, and sub-Saharan Africa, thereby evaluating the nuances of compliance and enforcement in diverse contexts. The research further endeavors to identify systemic vulnerabilities by investigating prominent data breaches, which illuminate the inherent weaknesses in information systems and cybersecurity infrastructures. Moreover, the study evaluates best practices and explores the potential of emerging technologies, including AI, ML, and blockchain, to mitigate these vulnerabilities and enhance data privacy. Finally, by anchoring its analysis in established theoretical frameworks, the study aspires to propose tailored solutions and policy recommendations that promote harmonized, yet locally adaptable, approaches to safeguarding sensitive patient information in an increasingly digitalized world.

Theoretical framework

The study draws on multiple theoretical perspectives to analyze the complex dynamics of healthcare data privacy, providing a structured foundation for understanding patient

trust, regulatory compliance, and the management of privacy among healthcare stakeholders. The rationale for integrating these theories is to capture both the human and technological dimensions of privacy challenges. Specifically, these theories help explain how trust is built and maintained, how innovations such as blockchain and other technologies can be adopted, and how privacy boundaries are negotiated in digital environments, a critical aspect given the multifaceted nature of healthcare data.

Social exchange theory, developed by George Homans and Peter Blau, highlights the critical role of trust in social interactions. Within healthcare data privacy, this theory underscores that patients are more likely to share sensitive health information when they are assured that their data will be protected.[26,27] In emphasizing trust as a currency in the patient–provider relationship, this theory supports our analysis of how ethical safeguards and transparent practices contribute to secure data sharing.

Everett Rogers's innovation diffusion theory provides insights into how new technologies and practices are adopted within social systems. Although first introduced in the early 1960s, its relevance persists today as it illuminates the processes by which privacy-enhancing technologies such as encryption, blockchain, and AI are embraced by healthcare organizations. This theory helps us understand barriers to innovation and suggests strategies for overcoming resistance, thereby facilitating the integration of advanced data protection measures.[28]

Travis Hirschi's social control theory emphasizes the importance of institutional and social mechanisms in deterring deviant behavior. In healthcare, this translates into the necessity for robust governance structures and organizational policies that ensure compliance with privacy regulations and prevent unauthorized access to sensitive patient data.[29] This theory underscores the need for regulatory oversight and internal controls in maintaining data integrity.

The ethical decision-making theory, influenced by scholars such as Lawrence Kohlberg, serves as a guide for navigating moral dilemmas in healthcare data privacy. This framework stresses the importance of balancing patient privacy rights with the imperatives of data sharing for research and treatment. It grounds our discussion in ethical principles of autonomy, beneficence, and justice, thereby guiding stakeholders in making informed decisions that respect patient rights while advancing healthcare objectives.[30,31]

Helen Nissenbaum's contextual integrity theory enriches the analysis by emphasizing the importance of maintaining privacy norms and expectations that are specific to social and professional contexts. This theory is particularly useful in explaining how healthcare settings, with their unique norms and relational dynamics, must adapt their data-handling practices to preserve confidentiality and trust.[32]

Finally, Elena Karahanna and Detmar Straub's information boundary theory explores how individuals establish and manage boundaries around their personal information in digital environments. This theory elucidates patient preferences regarding the control of access to their medical records and underscores the challenges of maintaining privacy in an era of pervasive digital connectivity.[33]

Together, these theories not only provide a comprehensive framework for analyzing the multifaceted challenges of healthcare data privacy but also directly inform our study's focus on the integration of emerging technologies and the management of privacy among healthcare stakeholders. In linking trust-building, innovation adoption, ethical decision-making, and privacy management, the theoretical framework underpins our recommendations for harmonized and adaptable data protection strategies.

Methodology

This study is a review that examines the multifaceted challenges and solutions associated with data privacy in healthcare. It employs a qualitative approach, integrating corpus construction and thematic analysis to explore global healthcare data privacy frameworks, identify vulnerabilities, and propose actionable recommendations.

Research design: corpus construction

The research adopts a corpus construction approach as a substitute for traditional sampling methods. Corpus construction involves selecting a set of relevant documents to form a representative body of knowledge on the subject matter. This approach is functionally equivalent to sampling but differs in structure, allowing the research to delve into various secondary data sources. The study constructed a robust corpus comprising regulatory documents, case studies, and scholarly articles. The corpus included key legal frameworks such as the GDPR in Europe, the HIPAA in North America, the APEC Privacy Framework, and sub-Saharan Africa's emerging data protection policies. These documents were sourced from reputable databases (e.g. PubMed, Scopus, Web of Science, Google Scholar, Wiley, and official organizational websites), official publications, and reports by regulatory authorities.[6,8,11,14] Case studies of significant health data breaches, such as the Anthem Inc. breach in the USA,[17] the WannaCry ransomware attack in the UK,[21] and the SingHealth breach in Singapore,[22] provided additional contextual insights.

As noted by Bauer and Aarts (2000, see Bauer and Aarts, Chapter 2 in this volume), "sample size does not matter in corpus construction as long as there is some evidence of saturation. Corpus construction is an iterative process where researchers continuously refine and expand their collection until no new insights emerge, thereby ensuring that the reliability and validity of the qualitative analysis are maintained." Furthermore, "in this sense, corpus construction and the representative sampling of textual and multimedia data enable researchers to capture the complexity of social

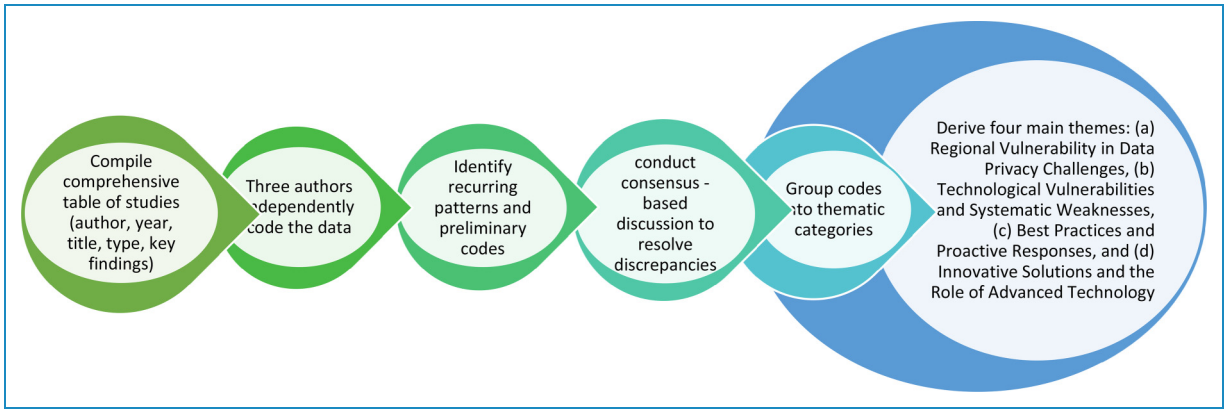

Figure 1. Coding process for deriving theme. This diagram illustrates the coding process: starting with the compilation of a comprehensive table of studies, the three authors independently coded the data to identify recurring patterns. Discrepancies were resolved through consensus-based discussions, leading to the grouping of codes into thematic categories and ultimately deriving four main themes (2025).

phenomena, thereby providing a robust foundation for qualitative analysis. In short, we contend that corpus construction typifies an iterative and reflexive process in which disparate data sources are systematically identified, selected, and organized into a coherent body of evidence. Secondly, we demonstrate corpus construction in the field by carefully curating documents from a range of reputable sources, ensuring that the resulting corpus accurately reflects the multifaceted dimensions of the research topic" (Bauer and Aarts, 2000).

To enhance clarity, a flowchart summarizing the methodological steps is provided in Figure 1.

Data sources

To ensure comprehensive coverage of the topic, we used databases such as PubMed, WHO databases, Africa CDC repositories, and government health ministries' websites. Additional reports and articles were sourced from PubMed, Scopus, Web of Science, Google Scholar, Wiley, and official organizational websites using the search terms: "healthcare data privacy," "regulatory frameworks," "data breaches in healthcare," "HIPAA," "GDPR," "APEC Privacy Framework," and "sub-Saharan Africa data protection."

Keywords were combined using Boolean operators (AND/OR) to ensure comprehensive retrieval of documents relevant to healthcare data privacy.

Inclusion and exclusion criteria

The inclusion and exclusion criteria were applied systematically during corpus construction to ensure that only documents meeting the study's relevance and quality thresholds were selected. To ensure rigor and relevance, this study employed clearly defined inclusion and exclusion criteria for selecting documents to be included in the analytical corpus. The inclusion criteria were as follows:

- Documents published between 2010 and 2024 to reflect both historical context and the most recent developments in healthcare data privacy.
- Peer-reviewed journal articles, policy briefs, legal frameworks, regulatory reports, and official communications from authoritative bodies such as the WHO, Africa CDC, European Data Protection Board (EDPB), and national data protection agencies.
- Case studies or reports involving notable healthcare data breaches or the implementation of major data protection regulations (e.g. GDPR, HIPAA, Protection of Personal Information Act (POPIA), APEC Privacy Framework).
- Materials addressing technical, regulatory, or ethical dimensions of data privacy within healthcare systems, including those focused on low-resource and high-resource contexts.

Documents were excluded if they:

- Were published before 2010, as these may not accurately represent current technological capabilities or regulatory environments.
- Provided only editorial opinions, news articles, or non-empirical content without substantive data or legal grounding.
- Focused exclusively on non-healthcare sectors or covered highly localized or outdated regulatory instruments with limited applicability to the broader study objectives.

These criteria were developed through iterative discussion among the research team, with consensus reached on ambiguous cases. The selected timeframe (2010–2024) was justified based on the proliferation of global data protection policies, the emergence of AI and digital health innovations, and the increasing frequency of high-impact data breaches during this period.

Data collection

Data collection followed a multipronged approach, emphasizing the triangulation of secondary sources. Legal texts, policy documents, and publicly available reports from organizations such as the EDPB,[9] the Office for Civil Rights,[7] and the Africa CDC[14] formed the backbone of the analysis. Peer-reviewed journal articles and industry reports supplemented these sources to provide a comprehensive perspective. This systematic approach ensures that the data collected is both comprehensive and directly relevant to the study's objectives.

Thematic analysis

A thematic analysis was conducted to systematically identify and extract recurring patterns from the corpus of selected documents. Initially, the research team compiled a comprehensive table summarizing each study by author(s), year, title, type of paper, and key findings. This table provided a structured foundation for our iterative coding process. During coding, emerging patterns were identified and grouped into distinct categories.

Through this process, four main themes emerged: (a) regional vulnerability in data privacy challenges; (b) technological vulnerabilities and systematic weaknesses; (c) best practices and proactive responses; and (d) innovative solutions and the role of advanced technology. These themes were derived inductively from the data and are directly aligned with the regulatory, technical, and strategic dimensions of healthcare data privacy identified in our review.

The coding was performed manually by three researchers, with any discrepancies resolved through consensus-based discussions. This iterative and collaborative approach ensured the transparency and rigor of the analysis, as well as a clear linkage between the emergent themes and the underlying data.

Addressing bias

The reliance on secondary data introduced the potential for selection bias. To mitigate this, the study incorporated diverse data sources, including regulatory texts, case studies, and academic research, ensuring a balanced representation of regional and global perspectives. Peer-reviewed articles were prioritized to ensure credibility, and the corpus construction and analysis were conducted independently by multiple researchers, with disagreements resolved through iterative consensus discussions. This approach enhanced the objectivity and reproducibility of the findings.

Comparative analysis

The study conducted a comparative analysis of healthcare data privacy regulations across North America, Europe, Asia, and sub-Saharan Africa. For example, North America's HIPAA Privacy Rule emphasizes legal obligations for protecting ePHI, while Europe's GDPR mandates comprehensive consent and accountability measures.[6,8] In contrast, the APEC Privacy Framework focuses on interoperability and voluntary cooperation, reflecting its member economies' diverse legal and cultural contexts.[11] Sub-Saharan Africa's policies, such as those endorsed by the Africa CDC, emphasize capacity building and incorporating cultural norms into privacy frameworks.[14] This comparative approach allows for a nuanced understanding of how different regulatory environments impact data privacy practices.

Case studies

Prominent case studies were analyzed to ground the findings in real-world scenarios. Examples include the Anthem Inc. breach, which exposed systemic weaknesses in cybersecurity and risk management in North America, and the WannaCry ransomware attack, which revealed vulnerabilities in outdated information technology (IT) systems in Europe. In the Asia-Pacific region, the SingHealth breach demonstrated the need for robust EHR systems. In contrast, sub-Saharan African examples, such as the Ghana Health Service's COVID-19 test results breach, underscored challenges related to encryption and governance.[17,21,22,24] These case studies were selected based on their ability to illustrate systemic vulnerabilities and inform the thematic analysis.

Ethical considerations

This study strictly adhered to ethical research practices appropriate for qualitative research involving document analysis and corpus construction. The selection and analysis of documents were conducted with a commitment to transparency, integrity, and accountability. All sources were obtained from reputable databases, official publications, and credible peer-reviewed journals, ensuring the validity and reliability of the findings. The collaborative approach among the authors further minimized potential biases.

Results

Through iterative coding and thematic analysis of the corpus, four primary themes emerged from the review of global healthcare data privacy frameworks:

Regional variability in data privacy challenges

The analysis revealed significant differences across regions:

- *North America:* Although the HIPAA Privacy Rule provides robust protections, enforcement challenges persist, as evidenced by breaches such as the Anthem Inc. and Premera Blue Cross incidents.

- *Europe:* The GDPR establishes strict data protection standards; however, events like the WannaCry ransomware attack on the UK's NHS expose vulnerabilities stemming from outdated IT systems and suboptimal patch management.
- *Asia-Pacific:* Integration of cross-border data privacy measures remains problematic. The SingHealth breach in Singapore, for example, underscores the gaps in data governance and the need for enhanced employee training.
- *Sub-Saharan Africa:* Resource limitations and inconsistent policy enforcement exacerbate vulnerabilities, with incidents such as the breaches in Ghana and South Africa highlighting region-specific challenges.

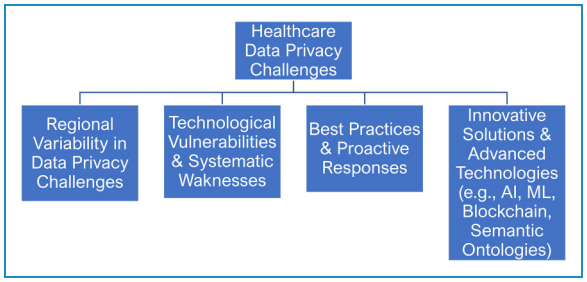

Figure 2. Thematic framework for global healthcare data privacy. This diagram illustrates the four main thematic areas derived from our review: (1) regional variability in data privacy challenges, (2) technological vulnerabilities and systemic weaknesses, (3) best practices and proactive responses, and (4) innovative solutions and advanced technologies.

Technological vulnerabilities and systemic weaknesses

Across all regions, there is a consistent pattern of systemic shortcomings, including outdated IT infrastructures, insufficient encryption practices, and gaps in cybersecurity protocols. These technological vulnerabilities contribute significantly to the risk of data breaches and impede effective data protection.

Best practices and proactive responses

Despite these challenges, several proactive strategies have been adopted:

- *North America:* In response to high-profile breaches, organizations have implemented comprehensive risk assessments and upgraded to advanced encryption protocols.
- *Europe:* The GDPR has catalyzed the development of disaster recovery plans and the institution of robust employee training programs, thereby strengthening data protection practices.
- *Asia-Pacific:* Regular cybersecurity audits, along with the implementation of stricter access controls, have been effective in mitigating risks, as seen in Singapore's response to the SingHealth breach.
- *Sub-Saharan Africa:* Initiatives by bodies such as the Africa CDC and the introduction of regulatory frameworks like POPIA have facilitated capacity-building and improved cybersecurity practices in the region.

Innovative solutions and the role of advanced technologies

The study also highlights the promising potential of emerging technologies to address existing gaps.

- AI and ML: these technologies involve algorithms and computational models that can learn from data, detect patterns, and make predictive decisions. In the context of healthcare data privacy, AI and ML can be leveraged to enhance breach detection systems, optimize data encryption algorithms, and automate compliance monitoring.
- Semantic ontologies: this refers to structured frameworks that define and relate concepts within a domain. In employing semantic ontologies, healthcare systems can improve data interoperability and understanding by establishing standardized relationships between data elements. This approach facilitates better data harmonization and supports more effective regulatory compliance and patient trust (Figure 2).

Discussion

The present study synthesizes a broad spectrum of literature on global healthcare data privacy, unveiling four interconnected thematic areas that not only reflect distinct challenges and opportunities but also provide avenues for policy evolution, technological innovation, and future research. In synthesizing over 90 studies (see Tables 1–5 and Tables A1–A5 in the Supplementary material), our analysis reveals significant regional disparities, systemic technological vulnerabilities, proactive regulatory responses, and the emerging role of advanced technologies in shaping this field. This comprehensive overview lays the groundwork for an in-depth discussion of how these themes inform current practices and future directions in healthcare data privacy.

Regional variability in data privacy challenges

The literature indicates that healthcare data privacy challenges differ significantly by region due to variations in regulatory frameworks, enforcement practices, and local

Table 1. Thematic framework for healthcare data privacy analysis.

Theme	Description and key examples
1. Regional variability in data privacy challenges	This theme highlights the significant differences in data privacy issues across regions. For instance, North America faces HIPAA enforcement challenges (e.g. Anthem Inc. breach), Europe grapples with vulnerabilities despite GDPR (e.g. WannaCry attack), the Asia-Pacific struggles with cross-border data governance (e.g. SingHealth breach), and sub-Saharan Africa contends with resource constraints and inconsistent policy enforcement (e.g. Ghana and South Africa breaches).
2. Technological vulnerabilities and systemic weaknesses	This theme encapsulates the systemic issues, such as outdated IT systems, insufficient encryption, and overall cybersecurity gaps that contribute to data breaches in healthcare across all regions.
3. Best practices and proactive responses	This theme outlines the proactive strategies adopted in different regions: North America has introduced enterprise-wide risk assessments and advanced encryption protocols; Europe has enhanced disaster recovery planning and employee training; Asia-Pacific has implemented regular cybersecurity audits and stricter access controls; and sub-Saharan Africa has seen capacity-building initiatives and regulatory improvements (e.g. POPIA).
4. Innovative solutions and the role of advanced technologies	This theme focuses on emerging technological solutions. It details how AI and ML can be used to improve breach detection, automate compliance monitoring, and optimize data encryption. Additionally, it covers the use of semantic ontologies to establish structured relationships between data elements, thereby improving interoperability and regulatory compliance.

Note. This table provides a visual summary of the thematic analysis derived from our review of global healthcare data privacy frameworks. Each theme is further discussed in the subsequent "Discussion" section, where its implications for policy and practice are elaborated upon.
HIPAA: Health Insurance Portability and Accountability Act; GDPR: General Data Protection Regulation; IT: information technology; POPIA: Protection of Personal Information Act; AI: artificial intelligence; ML: machine learning.

contexts. In North America, for instance, the enforcement challenges of HIPAA have led to high-profile breaches, as noted by Cohen[6] and exemplified in the Anthem incident.[39] In contrast, European healthcare systems benefit from the stringent requirements of the GDPR[9,10,40] but are not immune to systemic vulnerabilities, as evidenced by the WannaCry attack.[21] Similarly, studies from Asia-Pacific and sub-Saharan Africa[11–16] reveal that resource constraints and diverse legal interpretations further complicate the implementation of robust data protection measures. This regional heterogeneity suggests that while international frameworks provide a useful baseline, tailored approaches are necessary to address local challenges.[14,36] Such variability reinforces the need for policies that are both globally informed and locally adaptable.

Technological vulnerabilities and systemic weaknesses

A substantial portion of the literature underscores the technical deficiencies inherent in many healthcare systems. Outdated IT infrastructures, insufficient encryption methods, and gaps in cybersecurity measures contribute significantly to data breaches. For example, Mertens et al.[8] highlight how hidden data leaks can occur even in advanced systems, while Foreman et al.[42] and Goodday et al.[43] emphasize vulnerabilities in data capture and management. Furthermore, systematic reviews and case studies[17,37,44] consistently reveal that these technical shortcomings are compounded by systemic weaknesses, such as poor integration of new technologies and inadequate training of staff. The theoretical framework proposed by Nissenbaum[32] helps explain these contextual flaws by emphasizing the role of "contextual integrity" in maintaining privacy standards. Thus, addressing technological vulnerabilities requires a two-pronged strategy: immediate technical upgrades and long-term systemic reforms.

Best practices and proactive responses

In response to the multifaceted challenges of healthcare data privacy, a growing body of evidence highlights the significance of best practices and proactive regulatory interventions. Frameworks such as the GDPR,[49] the California Consumer Privacy Act, and South Africa's POPIA[51] have established rigorous benchmarks for data governance. These standards not only delineate legal obligations but also promote institutional accountability, secure data handling, and enforcement mechanisms that are increasingly viewed as models for global adoption.

Empirical studies from South Africa[15,16] illustrate how local implementation of such frameworks coupled with

Table 2. Regional variability in data privacy challenges.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
[1]	Marques and Ferreira (2020)	Digital transformation in health: a systematic review of 45 years of evolution	Systematic Review	Outlines the historical evolution and regional differences that impact digital health transformation.
[2]	Keshta and Odeh (2021)	Security and privacy of electronic health records: concerns and challenges	Review Article	Discusses challenges in EHR security that vary by region.
[3]	Fraser (2021)	Data privacy and security (in introduction to nursing informatics)	Book Chapter	Provides an overview of data privacy issues with regional nuances.
[4]	WHO (2024)	Health data privacy policy brief (online resource)	Policy Brief	Presents global health data privacy policies with regional implications.
[5]	UNESCO (2024)	Privacy policy (online resource)	Policy Document	Details international privacy guidelines influencing regional practices.
[6]	Cohen (n.d.)	HIPAA reform or a patchwork scheme	Policy/Review	Highlights enforcement challenges of HIPAA in North America.
[9]	Zarechahoki (2022)	How GDPR advances and harmonizes data contracts	Review Article	Demonstrates how GDPR sets benchmarks while revealing regional implementation challenges.
[10]	Granmar (2021)	Global applicability of the GDPR in context	Review Article	Explores GDPR's influence and regional discrepancies in its application.
[11]	Hiramatsu et al. (2021)	Current status, challenges, and future perspectives in Japan	Review/Policy	Reviews of real-world data challenges in Japan, highlighting regional practices.
[12]	Yasunaga (2020)	Protection of personal information in real-world data in Japan	Short Communication	Focuses on Japanese data protection practices in a regional context.
[13]	Miyashita (2021)	Human-centric data protection laws: a lesson from Japan	Review Article	Emphasizes region-specific legal approaches in Japan.
[14]	Daigle (2021)	Data protection laws in Africa: a pan-African survey	Survey/Review	Provides an overview of diverse data protection practices across Africa.
[15]	Staunton et al. (2021)	Data protection, management, and sharing in South Africa	Empirical Study	Explores stakeholder perspectives in South Africa's regulatory landscape.
[16]	Brand et al. (2022)	Data sharing governance in sub-Saharan Africa during emergencies	Review/Policy	Discusses governance challenges and regional variability in Africa.
[34]	Xiang and Cai (2021)	Privacy protection and secondary use of health data	Review Article	Outlines strategies that vary according to regional regulatory contexts.
[35]	Margam (2023)	Ethics and data privacy: the backbone of trustworthy healthcare	Commentary	Underlines ethical considerations that differ by region.
[36]	Tegegne et al. (2022)	Health professionals' knowledge of patient confidentiality	Cross-sectional Study	Examines regional attitudes and practices in resource-limited settings.

(continued)

Table 2. Continued.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
37	Oh et al. (2021)	A comprehensive survey on security and privacy for electronic health data	Survey Article	Provides comparative insights into regional technological and policy challenges.
38	Basil et al. (2022)	Health records database and inherent security concerns	Review Article	Reviews inherent vulnerabilities that are shaped by regional infrastructure.
39	DHHS, US (2020)	HHS breach portal notice	Official Report	Documents data breach cases in the USA, illustrating enforcement challenges.
40	Vretta (2018)	The new EU GDPR in medical data and clinical research	Commentary/ Policy	Discusses GDPR's influence as a model for regional data privacy practices.
41	Middelstorb (n.d.)	Balancing security objectives and data protection in the EU	Thesis	Analyses EU practices, providing a contrast with other regional approaches.

Note. This theme captures how different regions face distinct challenges in healthcare data privacy, reflecting variations in regulatory frameworks, enforcement practices, and contextual factors. This table illustrates the distinct challenges faced by different regions in healthcare data privacy. It highlights variations in regulatory frameworks, enforcement practices, and contextual factors that influence the protection of sensitive patient information. EHR: electronic health record; HIPAA: Health Insurance Portability and Accountability Act; GDPR: General Data Protection Regulation; NHS: National Health Service.

capacity-building and stakeholder engagement has led to measurable improvements in data privacy and cybersecurity infrastructure. Similarly, regulatory instruments in the USA, including the HIPAA[53] and the Gramm-Leach-Bliley Act,[52] have shaped privacy practices through stringent compliance requirements and post-breach corrective mechanisms.[17,21] These interventions underscore the value of regulatory foresight combined with technical resilience, particularly when deployed through cross-sectoral collaboration.

Across regions, proactive responses to breaches have become instrumental in evolving healthcare privacy norms. In North America, high-profile incidents such as the Anthem Inc. data breach prompted organizations to adopt comprehensive enterprise-wide risk assessments, implement advanced encryption standards, and provide identity theft protection services.[17,20] These efforts reflect a shift from reactive crisis management to preemptive risk mitigation and continuous system audits.

In Europe, the implementation of GDPR has catalyzed sector-wide changes, particularly following cyber incidents like the WannaCry ransomware attack. Institutions responded by strengthening disaster recovery protocols, investing in ongoing employee training, and deploying real-time vulnerability detection systems.[21] These practices have cultivated a culture of compliance and adaptability, where digital vigilance is integrated into operational workflows.

The Asia-Pacific region, guided by the APEC Privacy Framework, has emphasized interoperability and multilateral cooperation. In the aftermath of the SingHealth breach, Singapore undertook a comprehensive review of its EHR systems, adopting stricter access controls, enhanced audit trails, and regular cybersecurity audits.[22,23] These reforms illustrate the potential of combining regional policy alignment with technological innovation to build resilient health data ecosystems.

In sub-Saharan Africa, where resource limitations and fragmented policy enforcement present enduring obstacles, regional actors have pursued strategic responses. The Africa CDC, for example, has championed governance reforms, professional training, and regulatory harmonization across member states. Legislative instruments such as POPIA, along with targeted interventions in Ghana and South Africa, have advanced encryption use, fostered data protection awareness, and improved breach response mechanisms.[14,24,25] While challenges persist, these efforts demonstrate the efficacy of context-sensitive approaches that balance international standards with local capabilities.

Despite these advances, our review reveals persistent variability in implementation, technological readiness, and legal enforcement across jurisdictions. Discrepancies in the definition of sensitive health data, semantic inconsistencies in regulatory language, and infrastructural deficits continue to undermine standardization efforts. In response, there is a pressing need to develop interoperable frameworks and shared protocols for encryption, breach detection, and compliance auditing that are adaptable to both high-resource and low-resource settings.

Innovative technologies, including AI, ML, and semantic ontologies, offer transformative potential in addressing these gaps. AI and ML can automate compliance monitoring and enhance breach detection, while semantic ontologies support data harmonization by standardizing how information is interpreted across systems. Together, these

Table 3. Technological vulnerabilities and systemic weaknesses.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
[7]	Nelson (2024)	Navigating regulatory realities with blockchain in healthcare	Working paper/ preprint	Explore the blockchain's potential and current limitations in addressing technological vulnerabilities.
[8]	Mertens et al. (2023)	Google Tag Manager: hidden data leaks and EU data protection violations	Preprint/ technical report	Reveals how technical oversights can lead to significant data leaks.
[42]	Foreman et al. (2020)	Categorisation of adverse drug reactions in electronic health records	Research article	Highlights vulnerabilities in data capture and management systems.
[43]	Goodday et al. (2020)	Maximizing the use of social and behavioral data from EHRs	Research article	Discusses technical challenges in integrating complex data sources.
[37]	Oh et al. (2021)	A comprehensive survey on security and privacy for electronic health data	Survey article	Details systemic technological gaps affecting data security.
[38]	Basil et al. (2022)	Health records database and inherent security concerns	Review article	Identifies systemic weaknesses that predispose healthcare systems to breaches.
[44]	Silenou et al. (2021)	Digital health tools in Africa for pandemic control: scoping review	Scoping Review	Assesses technical readiness and vulnerabilities in African contexts.
[17]	Shankar and Mohammed (2020)	Surviving data breaches: a multiple case study analysis	Case study/ mixed methods	Illustrates real-world technical failures leading to breaches.
[18]	Flihan (2018)	Electronic crime in healthcare	Doctoral dissertation	Analyses technical dimensions of cybercrime in healthcare.
[19]	Roberts (2017)	Examining data breaches in healthcare	Doctoral dissertation	Provides a detailed investigation of systemic vulnerabilities.
[20]	Vedete (2015)	Methodologies of network defense	Doctoral dissertation	Presents technical defense strategies and their limitations.
[21]	Morse (2024)	Investigation: WannaCry cyber-attack and the NHS	Official report	Documents how outdated systems contributed to a significant breach.
[45]	Inkster et al. (2023)	Cybersecurity: a critical priority for digital mental health	Research article	Emphasizes the urgent need to address technical vulnerabilities in healthcare IT.
[22]	AlShkeili (2020)	Is the medical information of political elites at risk?	Research article	Examines targeted cyberattacks highlighting systemic tech gaps.
[23]	Lim (2021)	Cybersecurity and data protection: empirical observations in the PDPA context	Empirical study	Discusses technical challenges in implementing data protection measures.
[46]	Bangkok Post (2024)	The great data robbery	News article	Reports on high-profile data breaches due to technical failures.
[24]	Ghana Health Service (2024) (online)	COVID-19 test results exposure case	Web report	Illustrates technical vulnerabilities in health data management in Africa.

(continued)

Table 3. Continued.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
25	Van Niekerk (2017)	An analysis of cyber-incidents in South Africa	Research article	Provides evidence of systemic cybersecurity weaknesses in the region.
47	MG (2018)	Five massive data breaches affecting South Africans	News report	Highlights recurring technical vulnerabilities in healthcare IT systems.
32	Nissenbaum (2004)	Privacy as contextual integrity	Journal article	Introduces a theoretical framework to understand contextual flaws in data systems.
39	DHHS, US (2020)	HHS breach portal notice	Official report	Reinforces the prevalence of technical vulnerabilities through documented breaches.
2	Keshta and Odeh (2021)	Security and privacy of EHRs: concerns and challenges	Review article	(Also supports theme 1) Underlines technical shortcomings in securing EHRs.

Note. This theme focuses on technical gaps, outdated systems, and cybersecurity flaws that underlie many data breaches across healthcare systems. This table focuses on the technical gaps, outdated systems, and cybersecurity flaws that underlie many data breaches in healthcare systems. It summarizes the key findings related to the inherent vulnerabilities that compromise data security. IT: information technology; EHR: electronic health record.

tools contribute to scalable, context-responsive models of healthcare data protection that uphold both security and patient trust.

Drawing on insights from over 90 scholarly and policy documents, this analysis affirms that best practices in healthcare data privacy are most effective when they are proactive, technologically integrated, and regionally adaptable. Regulatory evolution, institutional investment in IT security, and cross-regional knowledge sharing will be critical to building secure and equitable digital health systems globally.

Innovative solutions and the role of advanced technologies

Emerging technologies offer promising avenues for overcoming current challenges in healthcare data privacy. The application of blockchain technology, as discussed by Nelson,[7] illustrates the potential for enhancing data integrity and transparency. Likewise, advances in AI and ML have been reviewed by Meng et al.[63] and Zhang et al.,[64] showing that these technologies can facilitate real-time breach detection, predictive analytics, and automated compliance monitoring. Semantic ontologies introduced by Tao et al.[67] and further developed by Marwadi[69] and do Espírito Santo and Medeiros[70] provide a framework for achieving improved data interoperability and standardization across heterogeneous datasets. Such innovative solutions not only address existing technical vulnerabilities but also enable a more integrated and adaptive approach to data privacy. The theoretical insights from Karahanna and Straub[33] on managing information boundaries further support the integration of these advanced technologies into existing systems.

Implications for research, policy, and practice

The findings of this review have direct and far-reaching implications for research, policy, and practice. For policymakers, the study emphasizes that effective data privacy management requires a multilayered, adaptive, and region-sensitive framework. Regulatory instruments must be grounded in globally accepted standards, such as GDPR, HIPAA, and POPIA, yet must remain flexible enough to accommodate local resource limitations and contextual differences. This approach ensures both international interoperability and contextual relevance.

Healthcare organizations must invest not only in technical enhancements—including encryption, blockchain, and AI—but also in operational reforms that prioritize continuous risk assessment, staff training, and internal governance. These investments should be viewed not merely as compliance obligations but as strategic imperatives for ensuring patient trust and institutional resilience.

For researchers, the review identifies a critical gap in evidence from underrepresented regions, notably Latin America and the Middle East. Future studies must incorporate primary data collection, particularly in low-resource settings, and assess the longitudinal impact of regulatory interventions and advanced technologies. In addition, exploring how AI, ML, and semantic ontologies can be

Table 4. Best practices and proactive responses.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
[15]	Staunton et al. (2021)	Data protection, management, and sharing in South Africa	Empirical study	Provides stakeholder perspectives on effective regulatory responses.
[16]	Brand et al. (2022)	Data sharing governance in sub-Saharan Africa during emergencies	Review/policy	Highlights governance strategies and capacity-building initiatives.
[48]	Cooperation AP (2005)	APEC Privacy Framework	Official document	Serves as a model for proactive cross-border collaboration on privacy.
[49]	GDPR (2018)	Regulation (EU) 2016/679	Legal/policy document	Sets high standards for data protection and serves as a best practice model.
[50]	CCPA (2018)	California Civil Code §§ 1798.100–1798.199	Legal/policy document	Provides guidelines that inform proactive privacy practices in the USA.
[51]	POPIA (2013)	Protection of Personal Information Act	Legal/policy document	Offers a framework for best practices in data protection in South Africa.
[50]	CCPA (2018)	California Civil Code §§ 1798.100–1798.199	Legal/policy document	(Duplicate) Reinforces proactive privacy measures through regulatory standards.
[52]	Federal Trade Commission (n.d.)	Gramm-Leach-Bliley Act	Legal/policy document	Outlines best practices for protecting consumer financial and health data.
[53]	HIPAA, US (n.d.)	Health Insurance Portability and Accountability Act	Legal/policy document	Establish baseline standards and corrective measures for data breaches.
[54]	PIPEDA, Canada (n.d.)	PIPEDA in brief	Legal/policy document	Demonstrates proactive regulatory measures in Canada.
[49]	GDPR (2018)	Regulation (EU) 2016/679	Legal/policy document	(Duplicate) Emphasizes the proactive role of GDPR in shaping global standards.
[55]	Va. Code (2018)	Virginia data privacy statutes	Legal/policy document	Provides regional legal standards that inform proactive responses.
[56]	National People's Congress (2021)	Personal Information Protection Law (China)	Legal/Policy document	Highlights China's proactive approach to data privacy regulation.
[57]	PDPA Singapore (n.d.)	Overview of the PDPA	Legal/policy document	Serves as a model for proactive privacy regulation in Singapore.
[58]	Data Protection Act, Kenya (2019)	The Data Protection Act	Legal/policy document	Sets forth best practices for data protection in Kenya.
[59]	Data Protection Act, Ghana (2012)	The Data Protection Act	Legal/policy document	Outlines regulatory measures in Ghana that serve as best practice examples.
[60]	Fortier et al., (2017)	Maelstrom Research Guidelines for retrospective data harmonisation	Guideline/ methodology	Provides rigorous guidelines that exemplify best practices in data integration.

(continued)

Table 4. Continued.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
61	Boland et al. (2017)	Ten simple rules to enable multisite collaborations through data sharing	Guidelines/ review	Provides practical guidelines for successful data-sharing collaborations.
17	Shankar and Mohammed (2020)	Surviving data breaches: a multiple case study analysis	Case study/ mixed methods	Describes corrective actions and risk assessments implemented post-breach.
21	Morse (2024)	Investigation: WannaCry cyber-attack and the NHS	Report	Details proactive responses (e.g. disaster recovery planning) following a significant breach.
45	Inkster et al. (2023)	Cybersecurity: a critical priority for digital mental health	Research article	Emphasizes the importance of and implementation of proactive cybersecurity measures.
62	Health Information and Quality Authority (2022)	[Title Not Specified]	Official report	Demonstrates institutional best practices in safeguarding health information.

Note. This theme outlines the proactive measures, regulatory frameworks, and policy-driven responses aimed at improving healthcare data privacy. This table outlines the proactive measures, regulatory frameworks, and policy-driven responses implemented to improve healthcare data privacy. It details best practices and governance strategies that mitigate privacy risks and enhance data protection. GDPR: General Data Protection Regulation.

integrated into legacy systems remains a key research frontier.

Policy recommendations

Effective data privacy management requires a multifaceted approach integrating technical, operational, and legislative measures. The following recommendations reflect both the universal standards and context-specific adaptations necessary for implementation.

Enhancing technical and operational requirements. Healthcare organizations should invest in advanced technologies like encryption, blockchain, and AI to secure patient data. Blockchain supports decentralized and tamper-proof data management, while AI strengthens breach detection and response capabilities. In resource-constrained settings, this may involve incremental adoption supported by donor agencies or international partnerships. Operationally, regular staff training, adoption of standardized data management protocols, and investment in secure digital infrastructure are critical to minimizing human error and improving accountability.

Establishing comprehensive legislative frameworks. Policies should align with global standards like GDPR, HIPAA, and POPIA while addressing regional realities. For instance, sub-Saharan Africa requires tailored legislative models that factor in limited digital infrastructure, enforcement capacity, and socio-economic diversity. Support from multilateral institutions could help build capacity for implementation and compliance. Harmonizing data privacy laws across jurisdictions will foster cross-border cooperation and mutual accountability.

Promoting governance structures and privacy awareness. Robust governance structures, including the institutionalization of data protection officers, are essential for ensuring compliance. Continuous education and awareness programs for all staff levels must be institutionalized, with particular attention to frontline health workers who often handle patient data. These initiatives are crucial in building a culture of privacy vigilance, especially in low-resource healthcare environments.

International collaboration and continuous adaptation. Policymakers and healthcare providers must work together to develop standardized protocols, share cross-border best practices, and foster international dialogue on evolving threats and technological solutions. This also includes regional cooperation in surveillance, compliance audits, and emergency response coordination. In adapting to emerging threats and innovations, healthcare systems must adopt iterative learning models that regularly update privacy protocols and risk frameworks.

To enhance clarity and guide policy implementation, Figure 3 illustrates the key domains that should inform healthcare data privacy policy development and execution.

Table 5. Innovative solutions and the role of advanced technologies.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
[7]	Nelson (2024)	Navigating regulatory realities with blockchain in healthcare	Working paper/ preprint	Investigate the potential of blockchain to enhance data integrity and security.
[8]	Mertens et al. (2023)	Google Tag Manager: hidden data leaks and their potential violations	Technical report/ preprint	Highlights vulnerabilities that can be mitigated through advanced tech solutions.
[63]	Meng et al. (2020)	A survey on machine learning for data fusion	Survey article	Reviews how AI and ML can be applied for more robust data integration and breach detection.
[64]	Zhang et al. (2018)	Multi-source heterogeneous data fusion	Conference proceedings	Explore techniques for integrating diverse data sources using AI.
[65]	Ounoughi and Yahia (2023)	Data fusion for ITS: a systematic literature review	Systematic review	Assesses methods for data fusion that can be adapted for health data interoperability.
[66]	Dong and Srivastava (2013)	Big data integration	Conference proceedings	Discusses approaches to integrating large datasets relevant for healthcare IT.
[67]	Tao et al. (2011)	Towards semantic-web-based representation and harmonisation	Conference proceedings	Introduces semantic ontologies to improve data interoperability and standardization.
[68]	Kawai et al. (2000)	Statistical approaches to accepting foreign clinical data	Research article	Presents statistical methods that support innovative data integration practices.
[69]	Marwadi (2002)	Ontological semantic integration model	Thesis	Proposes an ontological model that lays the groundwork for semantic interoperability.
[70]	do Espírito Santo and Medeiros (2017)	Semantic interoperability of clinical data	Conference proceedings	Examines frameworks for achieving semantic interoperability in healthcare.
[71]	Bleiholder and Naumann (2009)	Data fusion	Survey article	Reviews methodologies of data fusion, underpinning innovative integration approaches.
[72]	Jugel et al. (2016)	VDDA: automatic visualization-driven data aggregation	Journal article	Presents an automated approach to data aggregation that supports innovative analysis.
[73]	Nassiri et al. (2017)	Integrating XML and relational data	Conference proceedings	Discusses integration techniques critical for modern data management systems.
[74]	Cecchin et al. (2010)	XML data fusion	Conference proceedings	Explores XML-based data fusion methods that enhance interoperability.
[75]	Kumar et al. (2021)	Data harmonisation for heterogeneous datasets: a systematic review	Systematic review	Provides insights into harmonization methods that facilitate innovative data integration.

(continued)

Table 5. Continued.

Ref.	Author(s) and year	Title	Type of paper	Key findings/contribution
[76]	Porter et al. (2014)	Harmonisation and translation of crop modelling data to ensure interoperability	Research article	Although focused on crop modelling, it offers transferable methods for data harmonization.
[77]	Firnkorn et al. (2015)	A generic data harmonisation process for cross-linked research	Research article	Outlines a harmonization process that can be adapted to complex healthcare data systems.
[78]	Kim et al. (2021)	Trusted compliance enforcement framework for sharing big health data	Conference proceedings	Proposes an innovative framework that leverages advanced technologies for enhanced compliance.
[79]	Cuzzocrea and Damiani (2021)	Privacy-preserving big data exchange: models, issues, future directions	Conference proceedings	Reviews models for secure data exchange using cutting-edge technology.
[80]	Alshumrani et al. (2023)	A unified knowledge graph to permit interoperability of heterogenous digital evidence	Conference proceedings	Demonstrates how knowledge graphs can enhance semantic interoperability in data systems.
[81]	Gordo and Larlus (2017)	Beyond instance-level image retrieval: leveraging captions for semantic retrieval	Conference proceedings	Explores innovative methods for semantic retrieval that support advanced data integration.
[33]	Karahanna and Straub (2011)	Information boundary theory: where do we go from here?	Commentary/ review	Provides theoretical insights that inform innovative approaches to managing data boundaries.

Note. This theme focuses on emerging technological innovations, such as blockchain, AI, ML, and semantic ontologies, that promise to enhance data privacy through improved integration, automation, and interoperability. This table synthesizes emerging technological innovations, such as blockchain, artificial intelligence (AI), machine learning (ML), and semantic ontologies, which promise to enhance data privacy through improved integration, automation, and interoperability. It provides a comprehensive view of forward-thinking strategies to address current challenges in healthcare data privacy. IT: information technology; XML: extensible markup language.

Directions for future research

As data harmonization continues to evolve through the integration of new technologies, methodologies, and collaborative frameworks, the field is poised for transformational advancements. These emerging trends highlight both the progress made and the dynamic, adaptive nature of healthcare data integration in addressing complex data privacy challenges.

The ability to scale. With the exponential growth of health data, the scalability of harmonization procedures is becoming increasingly essential.[80] Organizations and institutions are now collecting data at unprecedented volumes and velocity, necessitating harmonization frameworks capable of managing vast, heterogeneous datasets without compromising efficiency or data quality. Future research should prioritize the development and testing of scalable, high-capacity harmonization systems that preserve performance across diverse platforms and settings.

Mechanization. Automation represents a promising avenue for reducing the manual labor traditionally required in data harmonization processes.[80] Tasks such as data matching, cleaning, and integration can be significantly accelerated through AI and ML algorithms. Research should explore the extent to which automation can enhance accuracy and efficiency in harmonizing disparate datasets while also identifying safeguards to mitigate automation-related biases or errors.

Combining new and emerging data sources. The integration of real-time data streams and Internet of Things devices introduces both opportunities and complexities for harmonization.[80] These novel data types require adaptive frameworks that can accommodate rapid data flows and maintain interoperability across devices and platforms. Future studies should investigate how harmonization strategies can be dynamically restructured to accommodate evolving data environments without sacrificing coherence or security.

Cutting-edge semantic and analytical technologies. Emerging semantic technologies and advanced analytics are reshaping how harmonized data can be interpreted and utilized. Techniques such as natural language processing and semantic ontologies play a crucial role in aligning the context and

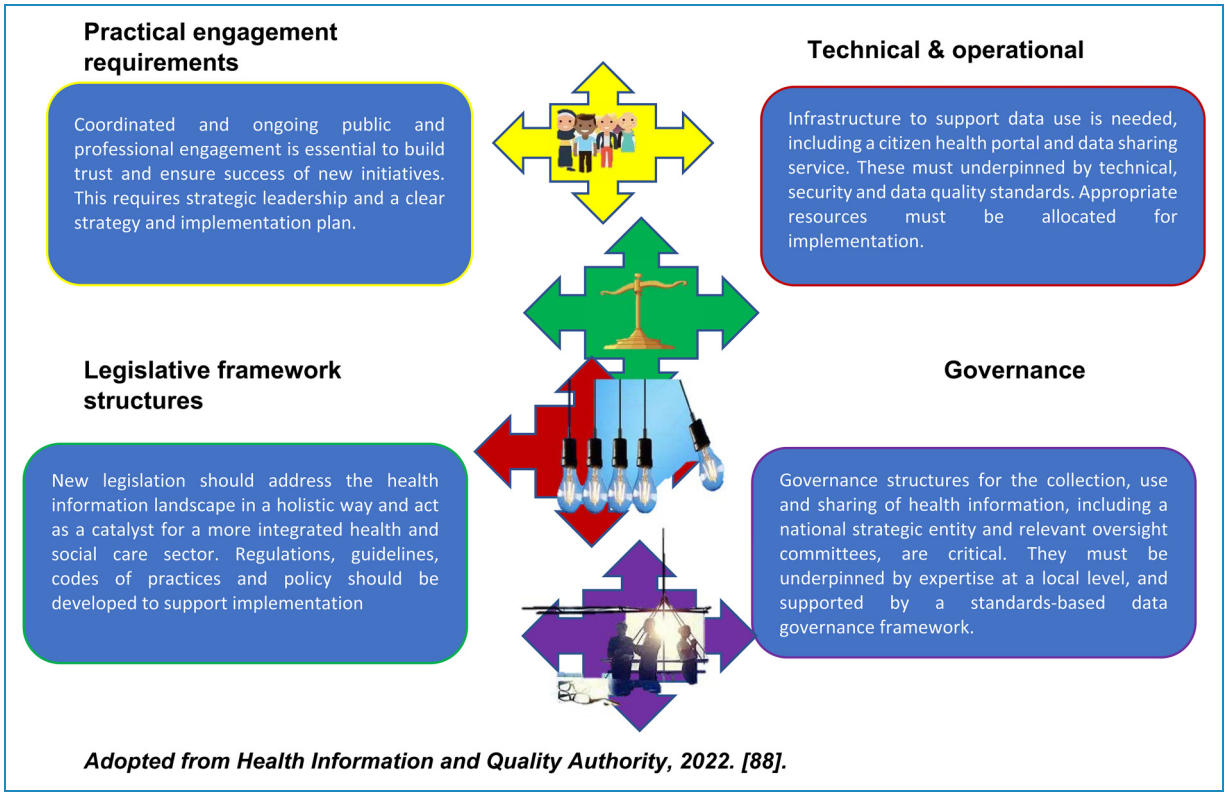

Figure 3. Key considerations to inform policy. Adopted from Health Information and Quality Authority, 2022.[62]

meaning of data from multiple sources.[61,77] Further research is needed to enhance these tools for greater semantic precision, allowing for more insightful, context-aware analyses, and reducing the risk of misinterpretation.

Real-time analysis and processing. There is a growing demand for real-time data processing and decision-making capabilities, especially in high-stakes sectors like healthcare and finance. The ability to assess and act on harmonized data in real time or near real time is becoming critical for timely clinical interventions, risk assessments, and operational efficiency. Research should focus on the development of frameworks that support real-time harmonization and analytics, including the necessary technical infrastructure and algorithms.[80]

Security, privacy, and ethical issues. As harmonization frameworks become more sophisticated, ensuring the security, privacy, and ethical use of sensitive data grows more complex and urgent. Future research must address the development of privacy-preserving methods and governance structures that comply with ethical standards and global regulatory requirements.[81] This includes exploring privacy-enhancing technologies, consent frameworks, and ethical auditing systems that ensure responsible data use while supporting integration and innovation.

Strengths and limitations of the study

This study is grounded in a robust review of 90 documents, including scholarly articles, regulatory frameworks, and case studies spanning four major world regions: North America, Europe, Asia-Pacific, and sub-Saharan Africa. The study's methodological strength lies in its systematic thematic analysis, which enabled the distillation of nuanced patterns and responses to healthcare data privacy across diverse contexts.

Additionally, the inclusion of detailed case studies (e.g. Anthem, WannaCry, SingHealth, and responses from Ghana and South Africa) provides empirical depth and contextual grounding for the thematic framework presented. The use of a rigorous coding process, a consolidated criteria for reporting qualitative research (COREQ) checklist, and visual aids (such as thematic diagrams) further enhances the transparency and reproducibility of this review.

However, the study acknowledges several limitations. First, the reliance on secondary data restricts the ability to capture emergent, localized dynamics that primary data could uncover. This includes real-time stakeholder experiences, evolving regulatory nuances, and the lived consequences of data breaches. Second, while the study covers all major global regions, it does not provide an in-depth analysis of Latin America and the Middle East, which limits the geographic generalizability of some conclusions.

Furthermore, although the literature selection was comprehensive, potential publication and language biases cannot be entirely ruled out. Some localized best practices or grey literature may not have been captured due to database constraints or language limitations. Finally, the fast-evolving nature of digital health technologies and privacy laws means that some of the findings may require periodic updating to remain current.

That said, the strength of this review lies in its methodological triangulation drawing from multiple case examples, thematic synthesis, and globally relevant regulatory frameworks to mitigate these limitations and produce generalizable, policy-relevant insights.

Conclusion

This review provides a critical synthesis of healthcare data privacy challenges and strategies across North America, Europe, Asia-Pacific, and sub-Saharan Africa. It highlights a global shift toward automated, intelligent data systems driven by regulatory reforms, technological innovation, and cross-border collaboration. While frameworks like GDPR, HIPAA, and POPIA offer strong legal foundations, disparities in implementation reveal the importance of adapting policies to local realities, especially in low-resource settings.

The findings emphasize that effective healthcare data protection requires multistakeholder engagement, combining government oversight, technological advancements such as AI and blockchain, and civil society participation. Although emerging technologies offer new opportunities for breach detection and interoperability, they also introduce ethical concerns regarding consent, autonomy, and data ownership that must be addressed.

Overall, this study underscores the need for an adaptive, equity-oriented approach to health data governance, balancing innovation with ethical safeguards. It offers practical, policy-relevant insights for strengthening global digital health systems, aligning with the WHO's vision for resilient, inclusive health governance in an interconnected world.

Acknowledgements

The corresponding author acknowledges the mentorship and quiet support received from academic colleagues whose encouragement made the completion of this manuscript possible.

ORCID iDs

Andrew Kweku Conduah https://orcid.org/0000-0001-6716-0939
Sebastian Ofoe https://orcid.org/0000-0002-2519-4222
Dorothy Siaw-Marfo https://orcid.org/0000-0001-8231-0755

Ethical approval

Not applicable. This study is based on publicly available secondary sources and did not involve direct interaction with human participants.

Consent to participate

Not applicable. No participants were recruited or surveyed for this review.

Author contributions

AKC led the thematic coding process and provided all financial resources. The emerging codes and categories were reviewed collaboratively with SHO, with discrepancies resolved through discussion. DS-M assisted with both technical formatting and the literature review. AKC took primary responsibility for the manuscript's content and revisions, with all authors approving the final version.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

The article and its Supplementary files include all relevant data for this study.

Reporting guidelines

Data Privacy in Healthcare: Challenges and Solutions, figshare dataset: https://doi.org/10.6084/m9.figshare.26532040.v1.

Supplemental material

Supplemental material for this article is available online.

References

1. Marques IC and Ferreira JJ. Digital transformation in health: A systematic review of 45 years of evolution. *Health Technol (Berl)* 2020; 10: 575–586.
2. Keshta I and Odeh A. Security and privacy of electronic health records: Concerns and challenges. *Egypt Inf J* 2021; 22: 177–183.
3. Fraser R. Data privacy and security. Introduction to nursing informatics 2021: 267–293.
4. World Health Organization (WHO). *Health data privacy policy brief*. Geneva: WHO, 2024. Available from https://www.who.int.
5. UNESCO. *Privacy policy*. Paris: UNESCO, 2024. Available from: https://www.unesco.org
6. Cohen D. HIPAA Reform or a patchwork scheme: A look at preemption, scope, and the inclusion of a Private Right of Action in a New Federal Data Privacy Law.
7. Nelson VR. Navigating regulatory realities with blockchain in healthcare: opportunities and limitations for telemedicine and electronic patient health records. Available at SSRN 4794825. 2024.

8. Mertens G, Bielova N, Roca V, et al. Google tag manager: hidden data leaks and its potential violations under EU Data Protection Law. arXiv preprint arXiv:2312.08806. 2023.
9. Zarechahoki A. How General Data Protection Regulation advances and harmonizes the international controller, processor, and data subject contracts. 2022.
10. Granmar CG. Global applicability of the GDPR in context. *Int Data Privacy Law* 2021; 11: 225–244.
11. Hiramatsu K, Barrett A and Miyata Y. PhRMA Japan medical affairs committee working group 1. Current status, challenges, and future perspectives of real-world data and real-world evidence in Japan. *Drugs – Real World Outcomes* 2021; 8: 459–480.
12. Yasunaga H. Protection of personal information in real-world data in Japan. *Ann Clin Epidemiol* 2020; 2: 1–2.
13. Miyashita Y. Ethical challenges of using health data for artificial intelligence in low-resource settings: a global health perspective. *Health Policy Technol* 2021; 10: 100534.
14. Daigle B. Data protection laws in Africa: a pan-African survey and noted trends. *J. Intl Com Econ* 2021; 1: 18–32.
15. Staunton C, Tschigg K and Sherman G. Data protection, data management, and data sharing: stakeholder perspectives on protecting personal health information in South Africa. *PLoS One* 2021; 16: e0260341.
16. Brand D, Singh JA, McKay AG, et al. Data sharing governance in sub-Saharan Africa during public health emergencies: gaps and guidance. *S Afr J Sci* 2022; 118: 1–6.
17. Shankar N and Mohammed Z. Surviving data breaches: a multiple case study analysis. *J Comp Int Manage* 2020; 23: 35–54.
18. Flihan NA. *Electronic crime in healthcare.* Doctoral Dissertation, Utica College, Utica, New York, USA, 2018.
19. Roberts SL. *Examining data breaches in healthcare.* Doctoral Dissertation, Utica College, Utica, New York, USA, 2017.
20. Vedete M. *Methodologies of network defense.* Doctoral Dissertation, Utica College, Utica, New York, USA, 2015.
21. Morse A. Investigation: WannaCry cyber-attack and the NHS. Report by the National Audit Office; 2024.
22. AlShkeili K. Is the medical information of political elites at increasing risk from strategically motivated cyberattacks? [unpublished manuscript]; 2020.
23. Lim S. Cybersecurity and data protection: some empirical observations, and a lacuna in the PDPA? [unpublished manuscript]; 2021.
24. Ghana Health Service. COVID-19 test results exposure case. Accra: Ghana Health Service; 2024 June 2. Available from https://ghs.gov.gh/.
25. Van Niekerk B. An analysis of cyber-incidents in South Africa. *Afr J Inf Commun* 2017; 20: 113–132.
26. Homans GC. Social behavior as exchange. *AJS* 1958; 63: 597–606.
27. Blau PM. *Exchange and power in social life*. New Brunswick, New Jersey, USA: Transaction Publishers, 1964.
28. Rogers EM. *Diffusion of innovations*. New York, USA: Free Press, 1962.
29. Hirschi T. *Causes of delinquency*. Berkeley: University of California Press, 1969.
30. Rest JR. Morality. In: Mussen PH (ed.) *Handbook of child psychology: vol. 3. Cognitive development*. 4th ed. Hoboken, New Jersey, USA: Wiley, 1983, pp. 556–629.
31. Kohlberg L. Stage and sequence: the cognitive-developmental approach to socialization. In: Goslin DA (ed.) *Handbook of socialization theory and research*. Chicago, Illinois, USA: Rand McNally, 1969, pp. 347–480.
32. Nissenbaum H. Privacy as contextual integrity. *Washington Law Rev* 2004; 79: 119–157.
33. Karahanna E and Straub D. Information boundary theory: where do we go from here? A commentary on the state of the field. *J Assoc Inf Syst* 2011; 12: 619–645.
34. Xiang D and Cai W. Privacy protection and secondary use of health data: strategies and methods. *BioMed Res Int* 2021; 2021: 1–11.
35. Margam R. Ethics and data privacy: the backbone of trustworthy healthcare practices. *Socio-Econ Human Aspects Township and Ind* 2023; 1: 232–236.
36. Tegegne MD, Melaku MS, Shimie AW, et al. Health professionals' knowledge and attitude towards patient confidentiality and associated factors in a resource-limited setting: a cross-sectional study. *BMC Med Ethics* 2022; 23: 26.
37. Oh SR, Seo YD, Lee E, et al. A comprehensive survey on security and privacy for electronic health data. *Int J Environ Res Public Health* 2021; 18: 9668.
38. Basil NN, Ambe S, Ekhator C, et al. Health records database and inherent security concerns: a review of the literature. *Cureus* 2022; 14: e23076.
39. US Department of Health and Human Services. Breach portal: Notice to the Secretary of HHS breach of unsecured protected health information. 2020. Available from: https://ocrportal.hhs.gov/ocr/breach (accessed 3 June 2024).
40. Vretta M. *The new EU General Data Protection Regulation (GDPR) in medical data and clinical research*. Master's Thesis, International Hellenic University, Thessaloniki, Greece, 2018.
41. Middelstorb M. *The European Union's balancing between security objectives and data protection: the case of passenger name record data*. Bachelor's Thesis. University of Twente, Enschede, Netherlands, 2018.
42. Foreman C, Smith WB, Caughey GE, et al. Categorization of adverse drug reactions in electronic health records. *Pharmacol Res Perspect* 2020; 8: e00550.
43. Goodday SM, Kormilitzin A, Vaci N, et al. Maximizing the use of social and behavioural information from secondary care mental health electronic health records. *J Biomed Inform* 2020; 107: 103429.
44. Silenou BC, Nyirenda JL, Zaghloul A, et al. Availability and suitability of digital health tools in Africa for pandemic control: scoping review and cluster analysis. *JMIR Public Health Surveill* 2021; 7: e30106.
45. Inkster B, Knibbs C and Bada M. Cybersecurity: a critical priority for digital mental health. *Front Digit Health* 2023; 5: 1242264.
46. Bangkok Post. *The great data robbery: Southeast Asia's digital heist*. Bangkok: Bangkok Post, 2024. Available from https://www.bangkokpost.com/.

47. MG. Five massive data breaches affecting South Africans. News Report, 2018.
48. Cooperation AP. *APEC privacy framework* (Vol. 81). Singapore: Asia Pacific Economic Cooperation Secretariat, 2005.
49. General Data Protection Regulation (GDPR). *Regulation (EU) 2016/679 of the European Parliament and the Council*. Available from https://eur-lex.europa.eu/eli/reg/2016/679/oj (2018, accessed 3 June 2024).
50. California Consumer Privacy Act (CCPA). *California Civil Code §§ 1798.100-1798.199*. Available from https://oag.ca.gov/privacy/ccpa (2018, accessed 3 June 2024).
51. Protection of Personal Information Act (POPIA). *Act No. 4 of 2013, Republic of South Africa*. Available from https://www.justice.gov.za/legislation/acts/2013-004.pdf (2013, accessed 3 June 2024).
52. Federal Trade Commission. *Gramm-Leach-Bliley Act*. Available from https://www.ftc.gov/business-guidance/privacy-security/gramm-leach-bliley-act (n.d., accessed 3 June 2024).
53. U.S. Department of Health & Human Services. *Health Insurance Portability and Accountability Act (HIPAA)*. Available from https://www.hhs.gov/hipaa/index.html (n.d., accessed 3 June 2024).
54. Office of the Privacy Commissioner of Canada. *PIPEDA in brief*. Available from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/ (n.d., accessed 3 June 2024).
55. Virginia Consumer Data Protection Act (VCDPA). *Va. Code Ann. § 59.1-571 - § 59.1-581*. Available from https://law.lis.virginia.gov/vacode/title59.1/chapter53/ (2021; accessed 3 June 2024).
56. National People's Congress. *Personal Information Protection Law of the People's Republic of China*. Available from https://www.chinalawtranslate.com/en/personal-information-protection-law-of-the-p-r-c/ (2021, accessed 3 June 2024).
57. Personal Data Protection Commission Singapore. *Overview of the PDPA*. Available from https://www.pdpc.gov.sg/Overview-of-PDPA (n.d., accessed 3 June 2024).
58. Data Protection Act, No. 24 of 2019 (Kenya). Available from https://www.odpc.go.ke/data-protection-act/ (accessed 3 June 2024).
59. Data Protection Act, 2012 (Ghana). Available from https://www.dataprotection.org.gh/data-protection-act (accessed 3 June 2024).
60. Fortier I, Raina P, Van den Heuvel ER, et al. Maelstrom research guidelines for rigorous retrospective data harmonisation. *Int J Epidemiol* 2017; 46: 103–105.
61. Boland MR, Karczewski KJ and Tatonetti NP. Ten simple rules to enable multi-site collaborations through data sharing. *PLoS Comput Biol* 2017; 13: e1005278.
62. Health Information and Quality Authority. *Guidelines on Information Governance*. Dublin: HIQA, 2022.
63. Meng T, Jing X, Yan Z, et al. A survey on machine learning for data fusion. *Inf Fusion* 2020; 57: 115–129.
64. Zhang L, Xie Y, Xiao L, et al. Multi-source heterogeneous data fusion. In: 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), 2018, pp. 47–51. IEEE
65. Ounoughi C and Yahia SB. Data fusion for ITS: a systematic literature review. *Inf Fusion* 2023; 89: 267–291.
66. Dong XL and Srivastava D. Big data integration. In: 2013 IEEE 29th International Conference on Data Engineering (ICDE), 2013, pp. 1245–1248. IEEE.
67. Tao C, Jiang G, Wei W, et al. Towards semantic-web-based representation and harmonisation of standard meta-data models for clinical studies. *AMIA Summit Transl Sci Proc* 2011; 2011: 59.
68. Kawai N, Andoh M, Uwoi T, et al. Statistical approaches to accepting foreign clinical data. *Drug Inf J* 2000; 34: 1265–1272.
69. Marwadi AK. *Ontological semantic integration model* [*dissertation*]. Kansas City, MO: University of Missouri-Kansas City, 2002.
70. do Espírito Santo JM and Medeiros CB. Semantic interoperability of clinical data. In: Data Integration in the Life Sciences: 12th International Conference, DILS 2017, Luxembourg, November 14–15, 2017, Proceedings; 2017, pp. 29–37. Cham: Springer.
71. Bleiholder J and Naumann F. Data fusion. *ACM Comput Surv* 2009; 41: 1–41.
72. Jugel U, Jerzak Z, Hackenbroich G, et al. VDDA: automatic visualization-driven data aggregation in relational databases. *VLDB J* 2016; 25: 53–77.
73. Nassiri H, Machkour M and Hachimi M. Integrating XML and relational data. *Procedia Comput Sci* 2017; 110: 422–427.
74. Cecchin F, de Aguiar Ciferri CD and Hara CS. XML Data fusion. In: Data Warehousing and Knowledge Discovery: 12th International Conference, DAWAK 2010, Bilbao, Spain, August/September 2010, 2010, pp. 297–308. Berlin: Springer.
75. Kumar G, Basri S, Imam AA, et al. Data harmonisation for heterogeneous datasets: a systematic literature review. *Appl Sci* 2021; 11: 8275.
76. Porter CH, Villalobos C, Holzworth D, et al. Harmonization and translation of crop modelling data to ensure interoperability. *Environ Model Softw* 2014; 62: 495–508.
77. Firnkorn D, Ganzinger M, Muley T, et al. A generic data harmonisation process for cross-linked research and network interaction. *Methods Inf Med* 2015; 54: 455–460.
78. Kim DY, Elluri L and Joshi KP. Trusted compliance enforcement framework for sharing big health data. In: 2021 IEEE International Conference on Big Data (Big Data), 2021, pp. 4715–4724. IEEE.
79. Cuzzocrea A and Damiani E. Privacy-preserving big data exchange: models, issues, future research directions. In: 2021 IEEE International Conference on Big Data (Big Data), 2021, pp. 5081–5084. IEEE.
80. Alshumrani A, Clarke N and Ghita B. A unified knowledge graph to permit interoperability of heterogeneous digital evidence. In: International Conference on Ubiquitous Security, 2023, pp. 420–435. Singapore: Springer Nature Singapore.
81. Gordo A and Larlus D. Beyond instance-level image retrieval: leveraging captions to learn a global visual representation for semantic retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017, pp. 6589–6598.