

Protocol

Developing an AI Governance Framework for Safe and Responsible AI in Health Care Organizations: Protocol for a Multimethod Study

Sam Freeman¹, PhD; Amy Wang¹, BSc, MPH; Sudeep Saraf², LLB, MBBS; Erica Potts², BSocSc(Psy); Amy McKimm², BASc; Enrico Coiera¹, MBBS, PhD; Farah Magrabi¹, BE, PhD

¹Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

²Alfred Health, Melbourne, Australia

Corresponding Author:

Sam Freeman, PhD

Australian Institute of Health Innovation

Macquarie University

Level 6, 75 Talavera Rd, North Ryde NSW 2109

Sydney

Australia

Phone: 61 2 9850 2400

Email: sam.freeman@mq.edu.au

Abstract

Background: Artificial intelligence (AI) has the potential to improve health care delivery through enhanced diagnostics, streamlined operations, and predictive analytics. However, health care organizations face substantial challenges in implementing AI safely and responsibly. This is due to regulatory complexity, ethical considerations, and a lack of practical governance frameworks. While many theoretical frameworks exist, few have been tested or adapted for real-world application in health care settings.

Objective: This study aims to develop and validate a practical AI governance framework to support the safe and responsible use of AI in health care organizations. The specific objectives are to identify governance requirements for AI in health care, examine existing AI governance processes and best practices, codevelop an AI governance framework to meet the needs of health care organizations, and test and refine the framework through real-world application.

Methods: A multimethod research design will be used, comprising four key stages: (1) a scoping review and document analysis to identify governance needs and current processes, (2) in-depth interviews with health care stakeholders as well as national and international AI governance experts, (3) development of a draft AI governance framework through a synthesis of findings, and (4) validation and refinement of the framework through stakeholder workshops and application to case studies of AI tools. Data will be analyzed using qualitative methods informed by grounded theory.

Results: The project received funding in October 2023. Ethics approval was obtained from the Alfred Health Human Research Ethics Committee (project 171/24) and the Macquarie University Human Research Ethics Committee (project 16508). Data collection commenced in April 2024, with the scoping review and document analysis being finalized. As of March 2025, a total of 43 interviews have been completed. The final AI governance framework is expected to be completed and ready for dissemination by June 2025.

Conclusions: This study will deliver a comprehensive AI governance framework co-designed with health care stakeholders to address real-world challenges in AI oversight. The framework will offer practical guidance to support health care organizations in adopting AI technologies safely, ethically, and in alignment with regulatory requirements. Outcomes from this study will inform local and international discussions on AI governance and promote the responsible integration of AI in health systems.

International Registered Report Identifier (IRRID): DERR1-10.2196/75702

(*JMIR Res Protoc* 2025;14:e75702) doi: [10.2196/75702](https://doi.org/10.2196/75702)

KEYWORDS

artificial intelligence; governance; health care; oversight; patient safety; ethics; framework

Introduction

Background

Artificial Intelligence (AI) technologies have the potential to transform health care by improving standards of care and addressing global challenges such as rising costs, staff shortages, and increasing patient complexity [1-4]. AI applications span both clinical and nonclinical tasks, including predicting patient care needs, assisting with diagnosis, and optimizing administrative processes [5]. Despite these benefits and technological advancements, adoption within health care organizations remains slow [6]. Compared to some other industries, health care is subject to significant regulation to ensure patient safety [7], and ethical considerations such as data privacy, transparency, and equity, which may limit adoption [8].

Governance at the organizational level is critical to ensure AI is implemented safely and responsibly. However, there is little practical guidance for governing AI development, implementation, and use in health care [9,10]. While numerous theoretical frameworks exist, these often focus on broad ethical principles such as fairness, transparency, and accountability but lack actionable strategies for health care organizations to adopt. This gap leaves organizations without clear mechanisms to assess AI risks, ensure compliance with regulatory standards, or integrate AI into clinical and operational workflows safely [8]. Most existing frameworks have not been designed for real-world application within complex health care environments. Few have been systematically tested or implemented, meaning there is limited evidence on their effectiveness in practice. Additionally, little is known about how key stakeholders such as health care professionals, patients, administrators, and AI developers perceive AI governance challenges or what governance mechanisms would best support responsible AI adoption [11]. Without a structured and adaptable governance framework, AI technologies may be deployed without sufficient oversight, leading to potential safety, ethical, and legal concerns.

Emerging evidence highlights safety concerns associated with AI and their impact on patient care, emphasizing the need for adequate governance to ensure a whole-of-system approach to safe AI implementation [12]. Effective governance at the organizational level is required to ensure that AI tools are effectively integrated into the IT infrastructure, as they are highly reliant on data from other clinical information systems [10]. For instance, data quality and requirements for any accompanying changes to the electronic medical record and other supporting clinical information systems need to be assessed to ensure data provided to the AI tools is fit for purpose and its output is accurately displayed to users. There is also a need to ensure that AI tools are integrated with the clinical workflow, with users aware of their intended use and trained to use them safely. Recent incidents of patient data being shared with the industry for AI training [13,14], along with evidence of hospitals deploying AI without adequate local evaluation for bias [15], also highlight the urgency for more stringent and consumer-centered AI governance.

The evolution of AI further necessitates a governance approach that can accommodate emerging technologies such as generative AI, AI-enabled clinical decision support systems, and automated administrative processes [16]. With the likelihood that AI capabilities will continue to progress, it is foreseeable that so will the potential risks, which will require governance approaches that are dynamic, iterative, and responsive to these technological and regulatory changes [17].

Health care must also consider organizational factors not addressed in broad conceptual frameworks, such as financial constraints, organizational culture, and commercialization [18]. Ongoing technological advancements and growing public awareness further necessitate governance strategies that adapt to emerging challenges. For example, the proliferation of generative AI, including AI scribes, has led to the development of new guidelines to address associated risks [19-21]. This study seeks to bridge these gaps by developing an AI governance framework tailored to the needs of health care delivery organizations, ensuring that AI innovations can be integrated safely and responsibly while maximizing their benefits for patient care and system efficiency [22].

Aim

This study aims to develop and test an AI governance framework to enhance organizational oversight of AI applications in health care delivery.

Objectives

To achieve the aim of the study, there are four specific objectives:

1. Identify governance requirements for AI applications in health care delivery organizations
2. Examine existing AI governance processes and current best practices for AI governance
3. Codevelop an AI governance framework to meet the needs of health care organizations
4. Test and refine the AI governance framework by applying it to exemplar AI tools

Output

We will produce a comprehensive AI governance framework for a health care delivery organization.

Methods

Study Setting

This study will be conducted within a major Australian public health care organization that delivers care to a catchment of over 770,000 residents in Melbourne's inner south. The organization operates across three hospital campuses, including a major tertiary and quaternary referral hospital that provides specialized care for patients who are critically ill. The other campuses focus on community-based health care, offering services in rehabilitation, geriatric medicine, general medicine, and aged mental health. Additionally, the organization delivers statewide and national health care services to patients across Victoria and Australia [23]. As a large and diverse health care provider, this setting offers a robust environment for developing

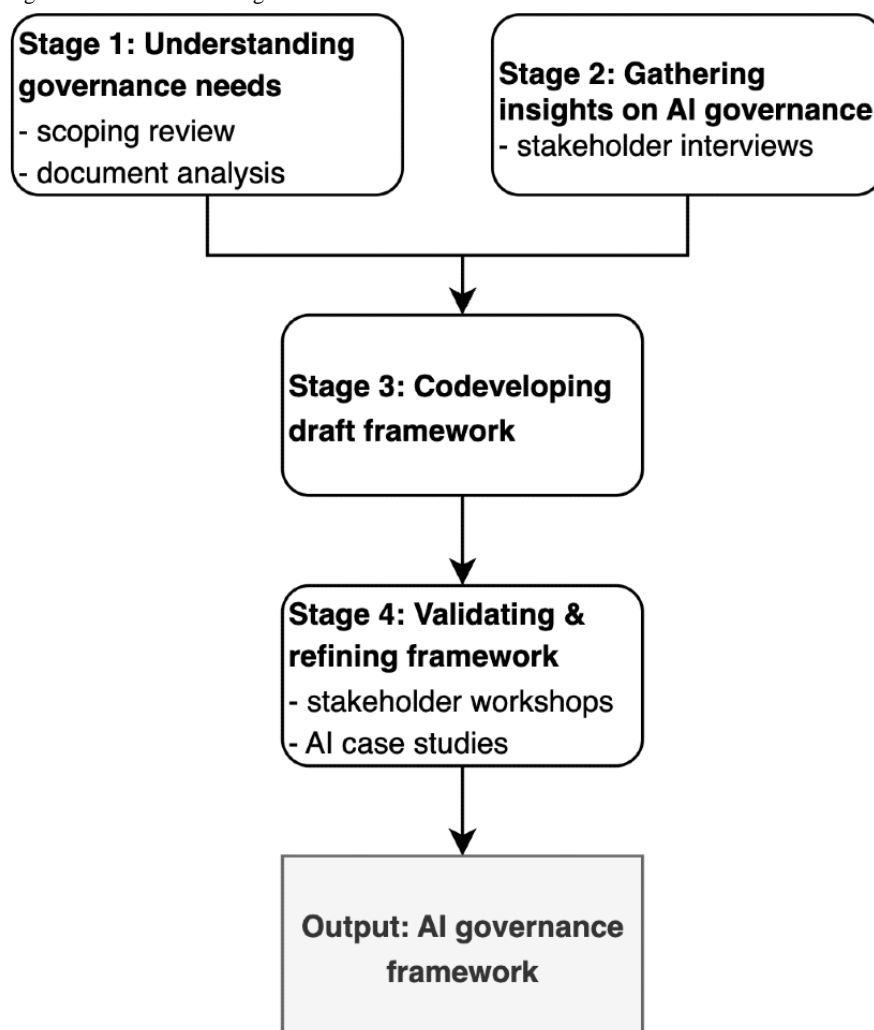
and testing an AI governance framework for real-world application in health care. While the organization already has established AI-driven initiatives and governance structures supporting patient safety and digital health, there is currently no standardized AI governance framework tailored to its varied operational needs.

Study Design

We will adopt an exploratory multimethod approach [24] to develop and test an AI governance framework for the safe and

responsible use of AI in health care delivery organizations. The structure of the study will comprise four stages that are directly linked to the study objectives: understanding governance needs through a scoping review and document analysis, gathering stakeholder insights on AI governance via in-depth interviews, codeveloping the AI governance framework through a synthesis of findings, and validating and refining the framework through stakeholder workshops and AI case studies (Figure 1).

Figure 1. Study process diagram. AI: artificial intelligence.



Given the evolving nature of AI governance within health care organizations, we have developed a comprehensive research strategy that draws on existing methodologies for investigating emerging technologies [25,26], evaluating complex systems [27], and ensuring stakeholder engagement in governance development [28]. This qualitative study will be designed in accordance with COREQ (Consolidated Criteria for Reporting Qualitative Research) [29]. The final output of the study is to produce a practical and actionable AI governance framework that addresses key challenges in AI oversight at the organizational level. The research will run concurrently to enable insights from one stage to inform other stages iteratively. The four research stages are detailed in the following sections.

Stage 1: Understanding Governance Needs

Scoping Review

Published literature about approaches to AI governance at an organizational level will be reviewed. Given the volume of literature published in this domain, only frameworks that are specific to health care will be reviewed [30]. The search strategy will be designed to address study site requirements, including governance approaches for AI tools with a clear intended purpose (eg, AI-enabled medical devices).

The following electronic databases will be searched: MEDLINE, Embase, and Scopus. These databases were selected based on consultation with a Macquarie University research librarian. A search will also be conducted on Google Scholar and Google

search, where the first 200 results will be included for review. This is based on guidance around the use of Google for searching for gray literature [31]. Reference lists of key papers will be hand-searched for additional peer-reviewed and gray literature. Appropriate search strings will be developed by linking concept clusters relating to each component of the question (AI, governance and implementation, and a health care setting) with Boolean operators “AND” and “OR.” Search results will be imported into Covidence where a two-stage review process (title and abstract screening, followed by full-text review) will be conducted by two independent reviewers. Data will be extracted into Excel (Version 16.98; Microsoft Corporation).

Document Analysis

To ensure the AI governance framework aligns with the current state of governance processes at the study sites, a document analysis will be conducted to map current processes. This will involve a systematic review of relevant internal policies, governance documents, and procedural guidelines. This will help establish a baseline of governance structures to ensure that the AI framework does not replicate existing or previously established governance processes and to identify strengths and gaps in current governance approaches.

Stage 2: Gathering Stakeholder Insights on AI Governance

Overview

In-depth semistructured interviews will be conducted with two key groups: health care staff and national and international experts in AI governance, policy, health care, AI, and patient safety. A purposive sampling technique will be used to identify key informants from both groups. A snowballing technique will also be used, allowing participants to refer other relevant informants for involvement in the study. Semistructured interview guides will be developed based on a review of the existing literature and input from the multidisciplinary project team. Members of the project team have recognized experience in health care, qualitative research, and AI safety and governance. Around 25-30 interviews (a total of 50-60) from both groups are expected to provide thematic saturation (where no new themes are identified). Key informants who agree to participate will be interviewed via online videoconferencing or in person, with audio recording. Audio files will be deidentified and professionally transcribed by a secure transcription service. Two investigators will independently analyze and code the interviews using open coding techniques based on grounded theory. A grounded theory approach will be applied as it provides a methodology and procedure for collecting and analyzing data and conceptualizing the results [32].

Health Care Stakeholder Interviews

We will recruit a diverse range of health care stakeholders to gain their perceptions on how to effectively govern the safe and responsible use of AI. The interviews will initially include informants from the board of directors and executive leadership, clinicians undertaking AI projects, clinical trial managers, and representatives from ethics and research governance offices, as well as clinical governance, risk, research, clinical operations,

medical services, hospital liaison, legal, and consumer representatives. During the interviews, participants will be asked about current approaches to governance and processes to manage the safe and responsible application of AI in health care and barriers to adoption. In addition to health care stakeholder interviews, dedicated consumer consultation will be incorporated into the framework's development process. To do this, we will engage the Australian Institute of Health Innovation Consumer Engagement Panel, which supports person-centered research. The panel will provide their perspectives on governance priorities through review sessions. Feedback from these sessions will be analyzed using a grounded theory coding approach and integrated alongside other stakeholder data to inform the development of the AI governance framework.

National and International Expert Interviews

We will examine current practices for governing AI in health care. Experts will include participants from academia (ethics, governance, digital health, technology, AI), government, health care administration, clinical settings, and professional health care associations. Participants will be asked about current approaches to AI governance in health care and will be invited to share any documentation about governing AI as well as tools or templates used to support AI governance processes.

Stage 3: Codeveloping a Draft AI Governance Framework

Findings from the scoping review, document analysis, and stakeholder interviews will be synthesized to develop a draft AI governance framework tailored to the needs of health care organizations. A grounded theory approach will guide the synthesis of findings across the scoping review, document analysis, and stakeholder interviews. Data from each source will be open coded to identify key concepts, which will then be compared and integrated through constant comparative analysis [33]. Emerging categories will be iteratively refined to develop the framework structure to reflect insights from stages 1 and 2.

The framework will outline pathways for integrating AI governance within existing clinical and operational workflows. To do this, the draft framework components will be mapped against current clinical governance, digital health, risk management, quality improvement processes, etc. Here, potential points of alignment and areas requiring modification will be identified. To address challenges such as organizational resistance, workflow incompatibilities, and training needs, the framework will include implementation recommendations outlining strategies for change management, staff education, and stakeholder engagement, informed by stakeholder insights.

During this phase, the framework will be designed to apply to different types of AI applications, including emerging technologies such as large language models and generative AI. A stratified governance approach based on the type of AI application and risk profile will be considered to address common and emerging risks, such as explainability challenges, hallucinations, etc, where identified. Additionally, specific governance components will be developed based on study findings to ensure the framework remains applicable across evolving AI technologies. The draft AI governance framework

developed in this stage will be tested and validated in stage 4 through stakeholder workshops and AI case studies.

Stage 4: Validating and Refining the Draft AI Governance Framework

Stakeholder Workshops

Stakeholder workshops will be conducted to refine and validate the AI governance framework with key stakeholders from the health care organization. Stakeholder workshops will be structured using facilitated discussion and scenario-based exercises focused on the practical application of the draft framework. The number of workshops and participants will be determined in consultation with the study site's project team to ensure appropriate representation of clinical, operational, and governance stakeholders. Detailed notes will be taken during the workshops, which will be analyzed using a grounded theory approach to identify key themes and areas for framework refinement. Insights from the workshops will be incorporated iteratively to improve the framework's practicality, ease of use, and alignment with organizational needs.

Testing the Draft Framework on Case Studies

The draft governance framework will be applied to exemplar AI tools from the study site to determine whether it provides adequate oversight. We will select 3-5 AI tools that are representative of the different pathways by which health care organizations acquire and use AI tools, including in-house development, codevelopment with external partners, and procurement of external AI tools. Selection criteria for case studies will also include representation of both clinical and administrative applications, diversity in levels of autonomy (eg, assistive vs decision support) [1], and variation in associated risk profiles. Tools will be selected in consultation with the study site's project team to ensure relevance and alignment with the organization's operational and clinical priorities. To evaluate the framework's usability and adaptability, a think aloud protocol will be used during pilot testing [34]. Participants will talk through their experience as they apply the framework to the AI tool in real time, providing insights into its effectiveness and identifying areas for improvement in the checklist. These findings will inform iterative refinement of the framework to enhance its applicability across diverse implementation contexts.

We will also assess the extent to which the AI framework will continue to be fit for purpose for current and next-generation AI technologies. Findings from these case studies and stages 1-3 will help identify gaps to refine the AI governance framework. The framework will be evaluated during this study based on its perceived usability, completeness, relevance to risk management and regulatory requirements, and stakeholder acceptance. This will be assessed with feedback collected during the stakeholder workshops and case study pilots.

Results

This study received funding in October 2023 from the Digital Health Cooperative Research Centre, Alfred Health, and Macquarie University. Ethics approval was granted by the Macquarie University Human Research Ethics Committee

(project 16508) in December 2023 and Alfred Health Human Research Ethics Committee (project 171/24) in February 2024. Data collection commenced in April 2024 and is progressing as planned.

As of March 2025, the scoping review and document analysis (stage 1) are underway, with preliminary findings identifying key themes in existing AI governance frameworks and internal governance structures at the study site. Interviews with health care stakeholders and AI governance experts (stage 2) are complete (n=43), and the thematic analysis has commenced. Case study selection for framework testing has been finalized in consultation with the study site. The final AI governance framework is anticipated to be finalized and disseminated by June 2025.

Discussion

Study Significance and Strengths

This study will generate critical insights into AI governance by developing and validating a framework for the safe and responsible integration of AI in health care. By addressing the gap in practical guidance for governance at the health care organizational level, this study will provide actionable strategies for managing AI tools in both research and operations. Its multimethod approach will ensure a comprehensive understanding of AI governance by integrating findings from a scoping review, stakeholder engagement, and case studies. This methodology will support the development of a governance framework that is both theoretically grounded and practically applicable.

A key strength of this study is the inclusion of diverse stakeholder perspectives, ensuring that governance recommendations are aligned with real-world requirements. By systematically developing an AI governance framework, this study will provide evidence-based recommendations for improving oversight of AI tools in health care delivery. It will ensure that organizational policy supports safe and responsible AI and complies with legislation (eg, data privacy, consumer law, and cybersecurity) and relevant AI ethics frameworks. The findings will contribute to international discussions on AI governance and support the development of governance that balances adaptability with patient safety and ethical considerations.

Limitations

Despite the strengths of this study, its limitations should be acknowledged. The study is currently limited to one Australian health care organization, which may limit the generalizability of the AI governance framework to other settings with different regulatory environments, organizational structures, or levels of technical maturity. While stakeholder interviews will include international participants, the findings may still reflect context-specific challenges rather than universal governance solutions. This also means that the successful integration of the framework into existing structures and organizational cultures may vary depending on local governance systems, staff engagement, and available training resources. Furthermore, while the framework is being developed at a large metropolitan

health care organization, limiting generalizability, it is being designed to be adaptable to support its application in diverse health care settings (eg, regional hospitals, resource-limited settings). However, further research will be needed to evaluate the framework's scalability and effectiveness across diverse health care environments.

Additionally, the study focuses on governance at the health care organizational level, which may not account for broader policy, legislative, or technological changes that might impact AI adoption in health care. The rapid progress of AI technologies such as generative AI presents challenges in ensuring the framework remains adaptable and relevant over time. While the iterative design of this study allows for ongoing updates, future research will be needed to assess the long-term sustainability and adaptability of the framework as AI capabilities continue to evolve.

Future Directions

This study will provide an AI governance framework specifically designed for health care organizations; however, further research will be needed to ensure its long-term applicability and scalability within dynamic health care contexts. Future studies should focus on implementing and evaluating the framework's effectiveness across a broader range of health care organizations, including regional and rural health services, to assess its adaptability in diverse settings. Expanding the

research to include international case studies will also provide comparative insights into how different regulatory and health care systems approach AI governance. Further research, including longitudinal follow-up studies, will be critical to track the sustained impact of AI governance frameworks on patient safety, operational efficiency, and clinician trust over time. This includes researching how governance processes develop alongside emerging AI technologies and regulatory changes as well as the long-term effectiveness and sustainability of the AI governance framework itself. This research should evaluate real-world outcomes such as long-term adoption, compliance improvement, and risk reduction.

Conclusions

This study will develop and validate an AI governance framework to support the safe and responsible use of AI in health care. By integrating a multimethod approach, the study will provide practical guidance on AI oversight at the organizational level, addressing gaps in existing governance structures. Findings will contribute to international discussions on AI governance, ensuring that AI technologies are adopted in a way that prioritizes patient safety, ethical considerations, and regulatory compliance. The iterative nature of this research will enable the framework to adapt to evolving AI capabilities and organizational needs, providing a foundation for robust governance strategies in health care.

Acknowledgments

The study was supported by the Digital Health CRC Limited, which is funded under the Australian Commonwealth's Cooperative Research Centres program, an Australian Government initiative for collaboration between university researchers and the industry.

Data Availability

The datasets generated during this study will not be made publicly available for privacy reasons, but requests to the corresponding author will be considered on a case-by-case basis.

Authors' Contributions

FM and EC conceptualized the study. FM, SF, and AW designed and are conducting the study. SF, AW, and FM drafted the manuscript with input from all authors. All authors provided revisions for intellectual content. All authors have approved the final manuscript.

Conflicts of Interest

SS and EP are employees of Alfred Health and participated in the project steering committee but were not involved in data collection, analysis, or interpretation. Their involvement in recruitment was limited to forwarding the study invitation to potential participants. SS participated in the board and chief executive officer interviews. SF, AW, FM, and EC have no professional or personal affiliation with Alfred Health.

References

1. Magrabi F, Lyell D, Coiera E. Automation in contemporary clinical information systems: a survey of AI in healthcare settings. *Yearb Med Inform*. Aug 2023;32(1):115-126. [FREE Full text] [doi: [10.1055/s-0043-1768733](https://doi.org/10.1055/s-0043-1768733)] [Medline: [38147855](https://pubmed.ncbi.nlm.nih.gov/38147855/)]
2. Lyell D, Coiera E, Chen J, Shah P, Magrabi F. How machine learning is embedded to support clinician decision making: an analysis of FDA-approved medical devices. *BMJ Health Care Inform*. Apr 2021;28(1):e100301. [FREE Full text] [doi: [10.1136/bmjhci-2020-100301](https://doi.org/10.1136/bmjhci-2020-100301)] [Medline: [33853863](https://pubmed.ncbi.nlm.nih.gov/33853863/)]
3. Coiera E, Liu S. Evidence synthesis, digital scribes, and translational challenges for artificial intelligence in healthcare. *Cell Rep Med*. Dec 20, 2022;3(12):100860. [FREE Full text] [doi: [10.1016/j.xcrm.2022.100860](https://doi.org/10.1016/j.xcrm.2022.100860)] [Medline: [36513071](https://pubmed.ncbi.nlm.nih.gov/36513071/)]
4. Jiang LY, Liu XC, Nejatian NP, Nasir-Moin M, Wang D, Abidin A, et al. Health system-scale language models are all-purpose prediction engines. *Nature*. Jul 2023;619(7969):357-362. [FREE Full text] [doi: [10.1038/s41586-023-06160-y](https://doi.org/10.1038/s41586-023-06160-y)] [Medline: [37286606](https://pubmed.ncbi.nlm.nih.gov/37286606/)]

5. Reddy S. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. *Implement Sci.* Mar 15, 2024;19(1):27. [FREE Full text] [doi: [10.1186/s13012-024-01357-9](https://doi.org/10.1186/s13012-024-01357-9)] [Medline: [38491544](https://pubmed.ncbi.nlm.nih.gov/38491544/)]
6. Han R, Acosta JN, Shakeri Z, Ioannidis JPA, Topol EJ, Rajpurkar P. Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. *Lancet Digit Health.* May 2024;6(5):e367-e373. [FREE Full text] [doi: [10.1016/S2589-7500\(24\)00047-5](https://doi.org/10.1016/S2589-7500(24)00047-5)] [Medline: [38670745](https://pubmed.ncbi.nlm.nih.gov/38670745/)]
7. Davenport TH, Glaser JP. Factors governing the adoption of artificial intelligence in healthcare providers. *Discov Health Syst.* 2022;1(1):4. [FREE Full text] [doi: [10.1007/s44250-022-00004-8](https://doi.org/10.1007/s44250-022-00004-8)] [Medline: [37521111](https://pubmed.ncbi.nlm.nih.gov/37521111/)]
8. Solanki P, Grundy J, Hussain W. Operationalising ethics in artificial intelligence for healthcare: a framework for AI developers. *AI Ethics.* Jul 19, 2022;3:223-240. [doi: [10.1007/s43681-022-00195-z](https://doi.org/10.1007/s43681-022-00195-z)]
9. Australia's AI ethics principles. Australian Government Department of Industry, Science and Resources. 2022. URL: <https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles> [accessed 2025-06-09]
10. Literature review and environmental scan final report: AI implementation in hospitals: legislation, policy, guidelines and principles, and evidence about quality and safety. Australian Commission on Safety Quality in Health Care. May 29, 2025. URL: https://www.safetyandquality.gov.au/sites/default/files/2024-08/artificial_intelligence_-_literature_review_and_environmental_scan.pdf [accessed 2025-06-09]
11. Embí PJ, Rhew DC, Peterson ED, Pencina MJ. Launching the Trustworthy and Responsible AI Network (TRAIN): a consortium to facilitate safe and effective AI adoption. *JAMA.* May 06, 2025;333(17):1481-1482. [doi: [10.1001/jama.2025.1331](https://doi.org/10.1001/jama.2025.1331)] [Medline: [39928518](https://pubmed.ncbi.nlm.nih.gov/39928518/)]
12. Lyell D, Wang Y, Coiera E, Magrabi F. More than algorithms: an analysis of safety events involving ML-enabled medical devices reported to the FDA. *J Am Med Inform Assoc.* Jun 20, 2023;30(7):1227-1236. [FREE Full text] [doi: [10.1093/jamia/ocad065](https://doi.org/10.1093/jamia/ocad065)] [Medline: [37071804](https://pubmed.ncbi.nlm.nih.gov/37071804/)]
13. Carter S, Prictor M. An imaging company gave its patients' X-rays and CT scans to an AI company. How did this happen? *The Conversation.* Dec 16, 2024. URL: <https://theconversation.com/an-imaging-company-gave-its-patients-x-rays-and-ct-scans-to-an-ai-company-how-did-this-happen-245753> [accessed 2025-06-30]
14. Powles J, Hodson H. Google DeepMind and healthcare in an age of algorithms. *Health Technol (Berl).* 2017;7(4):351-367. [FREE Full text] [doi: [10.1007/s12553-017-0179-1](https://doi.org/10.1007/s12553-017-0179-1)] [Medline: [29308344](https://pubmed.ncbi.nlm.nih.gov/29308344/)]
15. Nong P, Adler-Milstein J, Apathy NC, Holmgren AJ, Everson J. Current use and evaluation of artificial intelligence and predictive models in US hospitals. *Health Aff (Millwood).* Jan 2025;44(1):90-98. [doi: [10.1377/hlthaff.2024.00842](https://doi.org/10.1377/hlthaff.2024.00842)] [Medline: [39761454](https://pubmed.ncbi.nlm.nih.gov/39761454/)]
16. Esposito M, Palagiano F, Lenarduzzi V, Taibi D. On large language models in mission-critical IT governance: are we ready yet? *arXiv.* Preprint posted online on December 16, 2024. [doi: [10.48550/arXiv.2412.11698](https://doi.org/10.48550/arXiv.2412.11698)]
17. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. World Health Organization. Mar 25, 2025. URL: <https://iris.who.int/bitstream/handle/10665/375579/9789240084759-eng.pdf?sequence=1> [accessed 2025-06-09]
18. Alami H, Lehoux P, Papoutsis C, Shaw SE, Fleet R, Fortin J. Understanding the integration of artificial intelligence in healthcare organisations and systems through the NASSS framework: a qualitative study in a leading Canadian academic centre. *BMC Health Serv Res.* Jun 03, 2024;24(1):701. [FREE Full text] [doi: [10.1186/s12913-024-11112-x](https://doi.org/10.1186/s12913-024-11112-x)] [Medline: [38831298](https://pubmed.ncbi.nlm.nih.gov/38831298/)]
19. Artificial intelligence (AI) scribes: fact sheet. The Royal Australian College of General Practitioners. 2024. URL: <https://tinyurl.com/mrub7kxe> [accessed 2025-06-30]
20. Generative artificial intelligence and local governance models. Safer Care Victoria. 2023. URL: <https://tinyurl.com/4cxjnhb> [accessed 2025-06-09]
21. Meeting your professional obligations when using artificial intelligence in healthcare. Australian Health Practitioner Regulation Agency. 2024. URL: <https://www.ahpra.gov.au/Resources/Artificial-Intelligence-in-healthcare.aspx> [accessed 2025-06-30]
22. Baig MA, Almuhaizea MA, Alshehri J, Bazarbashi MS, Al-Shagathrh F. Urgent need for developing a framework for the governance of AI in healthcare. *Stud Health Technol Inform.* Jun 26, 2020;272:253-256. [doi: [10.3233/SHTI200542](https://doi.org/10.3233/SHTI200542)] [Medline: [32604649](https://pubmed.ncbi.nlm.nih.gov/32604649/)]
23. Annual report 2023-2024. Alfred Health. 2024. URL: <https://www.alfredhealth.org.au/about/governance/annual-report-2023-24> [accessed 2025-06-30]
24. Jupp V, editor. *The SAGE Dictionary of Social Research Methods.* London, United Kingdom. SAGE Publications, Ltd; 2006.
25. Crowe S, Cresswell K, Robertson A, Huby G, Avery A, Sheikh A. The case study approach. *BMC Med Res Methodol.* Jun 27, 2011;11:100. [FREE Full text] [doi: [10.1186/1471-2288-11-100](https://doi.org/10.1186/1471-2288-11-100)] [Medline: [21707982](https://pubmed.ncbi.nlm.nih.gov/21707982/)]
26. Gerring J. *Case Study Research: Principles and Practices.* 2nd ed. Cambridge, United Kingdom. Cambridge University Press; 2016.

27. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*. Sep 30, 2021;374:n2061. [FREE Full text] [doi: [10.1136/bmj.n2061](https://doi.org/10.1136/bmj.n2061)] [Medline: [34593508](https://pubmed.ncbi.nlm.nih.gov/34593508/)]
28. Helbig N, Dawes S, Dzhusupova Z, Klievink B, Mkude CG. Stakeholder engagement in policy development: observations and lessons from international experience. In: Janssen M, Wimmer MA, Deljoo A, editors. *Policy Practice and Digital Science: Integrating Complex Systems, Social Simulation and Public Administration in Policy Research*. Cham. Springer International Publishing; 2015:177-204.
29. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care*. Dec 2007;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
30. Batool A, Zowghi D, Bano M. AI governance: a systematic literature review. *AI Ethics*. Jan 14, 2025;5(3):3265-3279. [doi: [10.1007/s43681-024-00653-w](https://doi.org/10.1007/s43681-024-00653-w)]
31. Haddaway NR, Collins AM, Coughlin D, Kirk S. The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PLoS One*. 2015;10(9):e0138237. [doi: [10.1371/journal.pone.0138237](https://doi.org/10.1371/journal.pone.0138237)] [Medline: [26379270](https://pubmed.ncbi.nlm.nih.gov/26379270/)]
32. Halaweh M, Fidler C, McRobb S. Integrating the grounded theory method and case study research methodology within IS research: a possible 'Road Map'. 2008. Presented at: 7th IEEE/ACIS International Conference on Computer and Information Science; May 14-16, 2008; Portland, OR. URL: <https://aisel.aisnet.org/icis2008/165>
33. Hallberg LRM. The "core category" of grounded theory: making constant comparisons. *Int J Qualitative Stud Health Well-being*. Jul 12, 2009;1(3):141-148. [doi: [10.1080/17482620600858399](https://doi.org/10.1080/17482620600858399)]
34. Ericsson K, Simon H. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA. MIT Press; 1993.

Abbreviations

AI: artificial intelligence

COREQ: Consolidated Criteria for Reporting Qualitative Research

Edited by J Sarvestan; submitted 09.Apr.2025; peer-reviewed by A Adeoye, C Onah; comments to author 22.Apr.2025; revised version received 12.May.2025; accepted 28.May.2025; published 28.Jul.2025

Please cite as:

Freeman S, Wang A, Saraf S, Potts E, McKimm A, Coiera E, Magrabi F

Developing an AI Governance Framework for Safe and Responsible AI in Health Care Organizations: Protocol for a Multimethod Study

JMIR Res Protoc 2025;14:e75702

URL: <https://www.researchprotocols.org/2025/1/e75702>

doi: [10.2196/75702](https://doi.org/10.2196/75702)

PMID:

©Sam Freeman, Amy Wang, Sudeep Saraf, Erica Potts, Amy McKimm, Enrico Coiera, Farah Magrabi. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 28.Jul.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.