Article

A scoping review of the governance of federated learning in healthcare

Check for updates

Rebekah Eden [1,2] ✉, Ignatius Chukwudi[1,3], Chris Bain[4], Sebastiano Barbieri [1,5], Leonie Callaway[6,7], Susan de Jersey[6,7], Yasmeen George[4], Alain-Dominique Gorse [1,8], Michael Lawley [9], Peter Marendy[8], Steven M. McPhail[10], Anthony Nguyen [9], Mahnaz Samadbeik[1] & Clair Sullivan [1,7]

In healthcare, federated learning (FL) is emerging as a methodology to enable the analysis of large and disparate datasets while allowing custodians to retain sovereignty. While FL minimises data-sharing challenges, concerns surrounding ethics, privacy, maleficent use, and harm remain. These concerns can be managed by effective data governance. Data governance specifies procedural, relational, and structural mechanisms governing how data is captured, shared, and analysed, the resultant models and their use. However, limited insights exist on the optimal governance of this emerging technology. This study aims to develop a consolidated framework of the data governance mechanisms for FL in healthcare. A scoping review was performed, using deductive and inductive analysis of 39 articles. The framework includes twelve procedural, ten relational, and twelve structural mechanisms. The framework directs researchers to examine how to enact each mechanism and provides practitioners with insights into the mechanism to consider when governing FL.

To improve the sustainable and equitable delivery of healthcare, the healthcare industry needs to harness the potential of digital health data by rapidly learning from that data to transform healthcare practice and policy[1–3]. While the uptake of digital health technologies has resulted in a vast volume of digital health data, leveraging its potential has proven difficult due to data silos, privacy and security concerns, and governance and ethical challenges[4]. Federated learning (FL) is emerging as a promising technique that partially alleviates these challenges by enabling learning and adapting from large volumes of health data without sharing the data[5–7].

FL is a novel machine learning (ML) technique[8] that builds "AI (Artificial Intelligence) models on the basis of distributed and local datasets across multiple parties without data collection to prevent data leakage and privacy violation"[9]. In FL, unlike traditional ML, data remains within the data owners' infrastructure[10] and the model is sent to the data owner for training. The outputs of the analysis are then returned to the model provider, who continues this process with other data owners[9].

FL is revolutionising healthcare by enabling collaborative ML across institutions while preserving patient privacy[5,11,12]. In healthcare, FL has been used to predict heart-related hospitalisations[13], enhance brain MRI analysis[14,15], and improve early cancer diagnosis accuracy[16]; benefiting a range of diseases[17] without sharing patient data. Despite minimising data sharing challenges, FL in healthcare involves the secondary use of sensitive health data. This secondary use of data raises concerns about privacy, potential biases, and ethical issues. It also raises questions regarding what the secondary use of data entails and its potential for harm, including discrimination, stigmatisation, and unethical behaviour[18–20]. FL also introduces new concerns as it is susceptible to novel attacks, including model poisoning (e.g. malicious model attacks), model inversion (e.g. reverse engineering data), and model stealing (e.g. the data owner retains the model)[7]. Governance is required to manage these challenges[10].

Data governance provides a useful lens to understand how to govern FL in healthcare. Data governance examines the procedural (e.g. policies and procedures), relational (e.g. collaboration among stakeholders), and

[1]Queensland Digital Health Centre, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Herston, QLD, 4029, Australia. [2]UQ Business School, The University of Queensland, Blair Drive, St Lucia, QLD, 4072, Australia. [3]School of Information Systems, Queensland University of Technology, 2 George Street, Brisbane, QLD, 4000, Australia. [4]Faculty of IT, Monash University, Clayton Campus, Wellington Rd, Clayton, VIC, 3800, Australia. [5]Centre for Big Data Research in Health, University of New South Wales, Botany St, Sydney, NSW, 2052, Australia. [6]Centre for Health Services Research, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Herston, QLD, 4029, Australia. [7]Royal Brisbane and Women's Hospital, Metro North Health, Butterfield Street, Herston, QLD, 4029, Australia. [8]Queensland Cyber Infrastructure Foundation, 47 Staff House Rd, St Lucia, QLD, 4067, Australia. [9]The Australian e-Health Research Centre, CSIRO, 296 Herston Road, Herston, QLD, 4029, Australia. [10]Australian Centre for Health Services Innovation and Centre for Healthcare Transformation, School of Public Health and Social Work, Queensland University of Technology, Victoria Park Road, Kelvin Grove, QLD, 4059, Australia. ✉e-mail: r.eden@uq.edu.au

structural mechanisms (e.g. roles and responsibilities)[21] necessary to govern raw data and how it is captured, stored, shared, analysed, the algorithms developed, and how its outputs are used. Despite the importance of data governance for FL in healthcare, literature provides limited guidance and lacks a synthesised understanding of the necessary governance mechanisms. While literature provides some insights into the governance of traditional ML that remain relevant for FL, these are often in the form of broad principles (e.g., transparency, fairness, accountability) without detailed insights into how they are operationalised[22]. FL also poses unique infrastructure challenges resulting from distributed (rather than centralised) data amongst nodes[23,24], which complicates governance efforts[25,26]. Therefore, it is important to investigate, tailor, and establish governance mechanisms to meet the specific needs of FL.

The aim of this scoping review is to understand the governance mechanisms that can be used to underpin FL in healthcare. This involved: 1) identifying governance mechanisms from literature on FL governance in healthcare, and 2) identifying governance mechanisms from literature on the governance of related techniques (ML and federated data networks (FDN)) in healthcare that may be tailored for FL. These insights are leveraged to develop a consolidated account for the effective governance of FL in healthcare. This consolidated account provides a critical first step to form an evidence-base to guide healthcare teams when governing FL.

Results

As illustrated in Fig. 1, 39 papers were included in the review, which examined the governance of one or more techniques (i.e., FL, ML, FDN). Of the 39 papers (Supplementary Table 1), governing ML in healthcare was the predominant focus ($n = 31$), with limited research on governing FDN ($n = 5$) or FL ($n = 7$) in healthcare. Because articles can explore multiple techniques, the sum of the papers investigating the governance of ML, FL,

and FDN are greater than the total number of papers included. Most papers were conceptual, with only 12 empirical papers, half of which presented case descriptions of their organisation's efforts without reporting actual data.

Procedural mechanisms

Procedural mechanisms specify the parameters that guide the appropriate use of FL (and associated techniques) in healthcare. Four procedural governance subthemes which encompass twelve procedural mechanisms, were identified (Table 1): data privacy, formal guidelines and agreements, initial model utility, and ongoing monitoring. All procedural mechanisms were reported in ML studies. For FL and FDN, deterrence against misuse, initial evaluation, and model registration were not reported. The absence of studies investigating a mechanism for a specific technique does not suggest that the mechanism is irrelevant for the given technique, rather they have not been the focus of the studies included in the review.

In terms of data privacy, across all techniques (FL, ML, FDN, the need to protect the privacy of health consumers has been a core consideration. Concerns surrounding the potential for the reidentification[18,27,28] and misappropriation of health data[19,29] have been raised. This highlights the need for data access control mechanisms[29] and data de-identification[28]. In FL, despite data remaining within the data owner's infrastructure, the possibility exists for reidentification[5]. To mitigate this concern in FL, studies have suggested using synthetic datasets during initial training, encrypting data before learning, implementing differential privacy, authenticating access, and limiting the output provided to data users[5,10].

Across all techniques (FL, ML, FDN), studies have indicated that formal guidelines and agreements are required. Contractual agreements should be established between internal and external stakeholders before data provisioning to ensure data protection and intended use purposes are upheld[30]. In FL, contractual agreements are complex due to the sheer
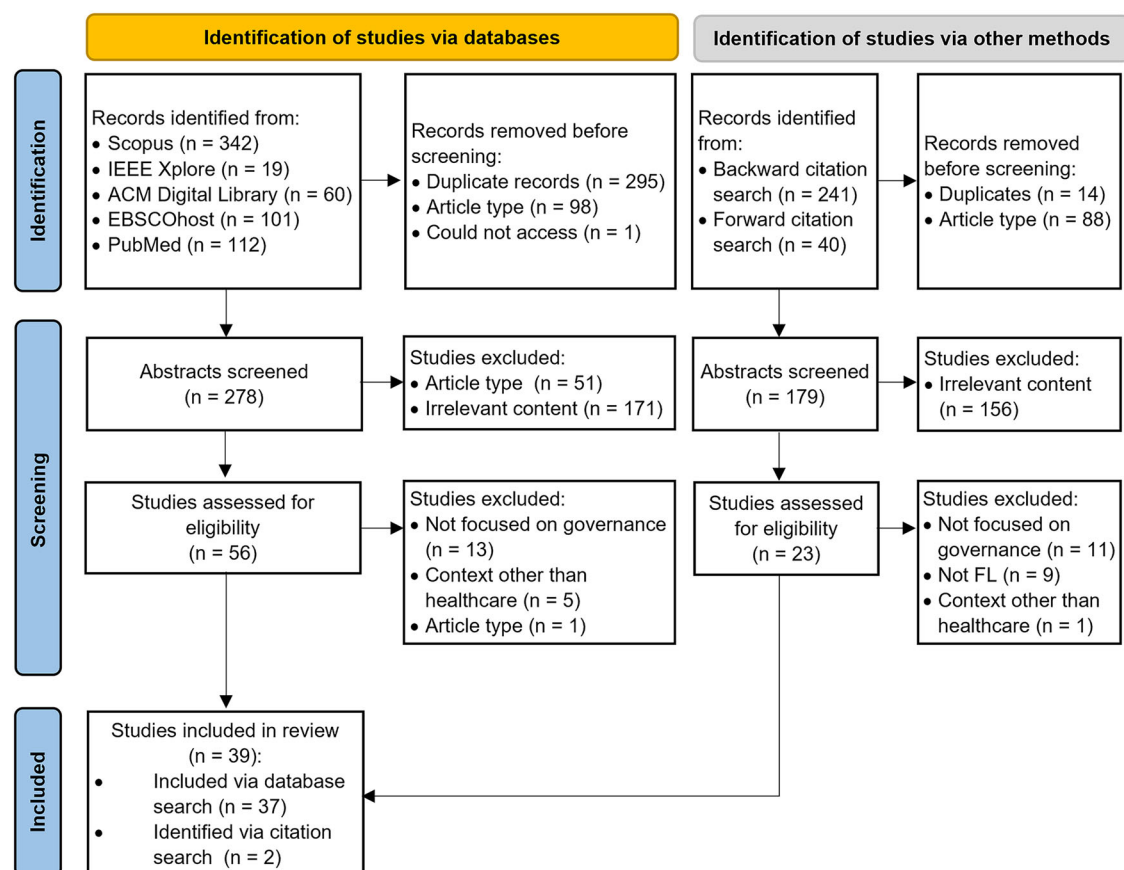

Fig. 1 | Article screening and selection process. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram detailing the article screening process.

Table 1 | Procedural mechanism theme

Subtheme	Mechanism	Description	FL	ML	FDN
Data privacy	Data deidentification	Approaches used to deidentify data for secondary use purposes	10,38	10,18,20,22,27,28,38,40,42,45,48	10
	Data provisioning and control	Techniques and processes used to regulate who has access to the data.	9,10,55	10,18,20,27–30,36,40,42,45,48	10,46,55
	Data security	The practices implemented to ensure data is kept secure and private.	5,10,32	10,29,35,36	10,46
Formal guidelines and agreements	Contractual agreement	Formal agreements between organizations regarding data sharing and use.	5,9,10,55	10,20,22,29,30,39,48,61	10,31,46,55
	Policy and procedures	"Guidelines and rules regarding the creation, acquisition, storage, security, quality, and permissible use of data."[21]	5,32,55	20,22,30,35,36,38,48–50,62–64	31,55
	Standards	Formalised data and operating standards guiding the collection, storage, sharing, and use of data.	5,10,55	10,20,22,28,30,34–36,62,65,66	10,31,46,55
Initial model utility	Initial evaluation	The evaluation of the efficacy and impact of the model prior to widespread adoption.		33–37,40,50,62	
	Model Registration	The register containing information surrounding the model, its characteristics, data, and intended use.		33,62	
Ongoing monitoring	Deterrence against misuse	Penalties for violating data security and access privileges.		19,27,36,38	
	Monitoring and auditing	Ongoing assessments of the extent to which the governance principles surrounding the data and its use are followed post-implementation.	5,9,10,32,55	10,18–20,22,27,29,33,34,37,39–41,45,49,50,62–67	10,46,55
	Risk management	Assessment and management of risks related to the development and use of models.	10	10,20,33,37,40,65	10
	Sustainability	Practices to ensure the longevity of the model.	5,10,55	19,39,62,64	10,51,55

Blank cells: Mechanism not investigated for the specific techniques in the include studies.
FL Federated Learning, *ML* Machine Learning, *FDN* Federated Data Networks.

number of parties involved and the need for any amendments to be synchronised[5]. Clear and formalised policies, procedures, and standards need to be developed and practised. In healthcare, these techniques (FL, ML, FDN) are often guided by FAIR (data is findable, accessible, interoperable, reusable) principles[10] and ideally leverage common data models (e.g., OMOP (Observational Medical Outcomes Partnership))[31] and interoperability standards (e.g., HL7 FHIR (Fast Healthcare Interoperability Resources))[28], which facilitate the secondary use of data. Using commonly agreed upon policies, procedures, and data standards is important in FL as it helps ensure consistent data pre-processing across nodes, thereby enhancing the integrity and reliability of the FL model[32].

Regarding initial model utility, ML studies have reported the need for ML models to be registered[33] and independently evaluated[34] before their use. Model registration involves recording details related to the model's data, characteristics, and intended use to improve transparency[33]. Initial model evaluation measures the model's efficacy with clinical data against pre-established performance indicators[33,35]. This is necessary to facilitate trust and ensure that the model will not be used in a discriminatory manner[36].

The potential for harm from using these techniques (FL, ML, FDN) has necessitated ongoing monitoring and risk management in their development and use[37]. Ongoing monitoring is needed to detect security breaches, access privilege violations, and misappropriation[19,36]. ML studies have also reported the need to deter ML model misuse, with financial penalties to prevent maleficent actors from unauthorised and discriminatory use of ML[27,38]. Monitoring is also needed to continuously assess the efficacy of ML, with refinements[39] and maintenance[19] necessary. Across all techniques (FL, ML, FDN), the sustainability of the predictive models needs to be actively managed. In FL, sustainability considerations are particularly important as reproducibility can be hampered if a data owner withdraws support[10].

Relational mechanisms

Relational mechanisms shape the interactions between the stakeholders implicated by FL (and associated techniques) in healthcare. Four relational governance subthemes, which encompass ten mechanisms, were identified (Table 2): capability, ethics, involvement, and institutional support. All were reported as important considerations of FL, ML, and FDN.

Across all techniques (FL, ML, FDN), studies reported the need for ethics and consent procedures. Ethical considerations emphasise that "new uses of people's data can involve both personal and social harms, but so does failing to harness the enormous power of data"[39]. These techniques should be underpinned by ethical principles that guide clinical care, medical research, and public health, such as respect for autonomy, equity, transparency, beneficence, accountability, and non-maleficence[37,40,41]. Given that a central tenet of ethics is informed consent[41], individuals should know how their data is used, by whom, and any commercial benefits[40]. This is also applicable in FL, as despite data not being shared, ethical considerations remain surrounding how the data is used and by whom[10]. However, obtaining informed consent is considered impractical due to the large volumes of health data used[42]. This has also led to discussions surrounding opt-in versus opt-out consent[39]. In addition, individual consent is not necessarily a requirement for ethical use of data in circumstances where potential beneficence outweighs risk in light of appropriate protections. In these cases, gatekeeper consent from data custodians who have weighed ethical considerations has been the norm, and may also become the norm for FL[10].

The importance of stakeholder involvement was also reported for all techniques (FL, ML, FDN). Although de-identified data reduces the requirement for informed consent[22], the public needs to be informed and accepting of how their data will be shared and used[37]. This requires strong consumer involvement and communication with community juries to foster the development of a social licence[27,43]. If health data is used in ways, or by actors, that are at odds with the interests of health consumers[10], the consumer social licence would be violated[10]. In addition to health consumers, strong engagement amongst all stakeholders (e.g., healthcare organisations, clinicians, regulators, developers, researchers, public/private organisations, vendors) is required[35,37]. In FL, stakeholder engagement is

necessary to establish a shared understanding of the vision and objectives of the project amongst all nodes[10]. This requires significant coordination between the nodes[32] and negotiation regarding decentralised and centralised infrastructure provisioning costs[10]. Consumer involvement and broader stakeholder engagement are necessary to engender trust[44].

Meaningful stakeholder engagement and clinical involvement will require capability development amongst all stakeholders implicated in digital health data governance, including clinicians[37] and health consumers[40]. Education and training are needed to improve data, algorithm, and digital literacy. This will enable clinicians to act as an "intermediary between developers and regulators"[37] and to understand how to interpret and act upon insights from AI technologies in their day-to-day work[35]. This will also enable health consumers to make informed decisions and meaningfully shape ML initiatives[29,40,45]. In FL, significant efforts are needed to develop the capabilities of the data stewards within each node, which will involve training sessions paired with auxiliary documentation regarding pre-processing and monitoring activities[32].

Studies report that institutional support is necessary for creating an environment conducive to using all techniques (FL, ML, FDN), including cultural management, leadership, and financial provisions. These techniques require a cultural shift involving the development and maintenance of cultural values such as trust, transparency, learning, and accountability in the use of data[19,35]. In FL, cultural differences between stakeholders must be adequately managed (e.g., commercialisation versus open science)[10]. Strong leadership and a vision in which AI is positioned as foundational to underpinning improvements in health and care are required for these techniques to succeed[10,41]. Financial considerations are also necessary, requiring sustainable business models[46] and support from funding bodies[10]. In FL, debates have been raised regarding whether financial incentives should be provided by the FL model providers to the data owners[9].

Structural mechanisms

Structural mechanisms specify the roles and responsibilities necessitated by FL (and associated techniques) in healthcare. Three structural governance subthemes, which encompass twelve structural mechanisms, were identified (Table 3): establishing oversight bodies, establishing roles, and establishing and considering health consumers. The term health consumer rather than patient is used to denote "anyone who has used, currently uses, or will use health care services …[and] represent[s] the person's more active role in making healthcare and medical decisions with their clinicians"[47]. As evident in Table 6, there was variation in how these were considered by studies investigating FL, ML, and FDN.

Oversight bodies include ethical boards, advisory boards, notified bodies, regulatory boards, and publication review groups that develop and implement safeguards surrounding the use of health data need to be established. Ethical boards, including human research ethics committees and institutional review boards, oversee the ethical conduct of research[18,27]. In the context of FL, questions remain about where the ethical board should be situated[10]. In traditional ML, ethical approval and oversight are typically sought from the data user's institution[10]. In FL, as data is not shared and the analysis is performed within the data owners' infrastructure, it may be appropriate for ethical approval to be sought from the data owner's institution, which is the responsibility of data stewards at each node[10]. Advisory boards should have diverse membership across health, legal, and security domains, as well as health consumer advocacy groups, to meaningfully oversee and provide guidance related to the use of data[31,48]. Notified bodies are delegated responsibility from regulatory bodies to audit and approve ML-equipped medical devices before widespread adoption[49]. Publication review groups need to be established[48] to ensure publications resulting from the use of data are consistent with ethical considerations, policies, and procedures. In FL, publishers need to encourage transparent predictive model sharing rather than data sharing[9].

There are diverse roles across FL, ML and FDN, which need to be established. These include data-safeguarding entities, developers, and project management teams. Data safeguarding entities include data owners,

Table 2 | Relational mechanism theme

Subtheme	Mechanism	Description	FL	ML	FDN
Capability	Capability development	Provision of training and education programs to improve digital literacy, data literacy, and algorithmic literacy of stakeholders and consumers.	10,32	19,20,22,27,29,30,35,37,40,45,50,63,66,67	46
Ethics	Ethical considerations	Consideration and adherence to the ethical principles and values governing the collection, storage, and use of data.	10	10,18,20,22,27–29,34,37,39–42,45,48,49,64–67	10,31,46
	Consent	Modes in which permission is sought and provided by health consumers for the secondary use of data.	10	18–20,22,27–29,34–37,39–43,45,48,61,62,66	10,46
Involvement	Consumer involvement	Active involvement of individuals in decisions surrounding the collection, storage, and use of their data.	10	10,18,20,22,27,29,37,39–41,43–45,48,49,66	10
	Establishing trust	Establishment of trust with consumers and between collaborating partners during data collection, sharing, and use.	5,10	9,10,27,29,35,36,39,42,44,45,49,66	10
	Stakeholder communication	Communication, collaboration, and coordination between a diverse array of stakeholder groups.	5,9,32	19,20,29,30,35,36,39,40,50,62,66	31
	Stakeholder engagement	Active involvement of a diverse array of stakeholders other than consumers during the AI initiative.	10,32,55	19,20,22,30,34–38,41,42,44,45,49,50,62,65–67	10,46,55
Institutional Support	Culture management	Developing and maintaining a culture that is values trust, transparency, and accountability in the use of data.	10	10,19,33,35,40,48,49,65	10
	Financial provisions	Allocation of financial resources to support the provision of sociotechnical resources for the sustainability of AI.	5,9,10,55	27,45,50	46,55
	Leadership	The vision and commitment of leaders to leverage AI models and provision resources to support the vision.	10	20,41,48,65	51

Blank cells: Mechanism not investigated for the specific techniques in the include studies.
FL Federated Learning, *ML* Machine Learning, *FDN* Federated Data Networks.

Table 3 | Structural mechanism theme

Subtheme	Mechanism	Description	FL	ML	FDN
Establishing oversight bodies	Advisory board	The body consisting of healthcare providers, healthcare stakeholders, and/or technical experts that are responsible for providing oversight over the use of health data.		18,20,33–35,41,48,49,64	31
	Ethical board	The entity responsible for overseeing and approving the ethical conduct of research.	10,32	10,18,27,42,67	10,51
	Notified body	The entity responsible for performing conformance assessments and ensuring compliance with requirements[9].	46	49,62,66	46
	Publication review group	The entity responsible for ensuring publications derived from data sharing comply with participant consent and adhere to policy.	10	30,48	
	Regulatory Body	The entity responsible for ensuring medical devices and systems leveraging AI are safe and effective.		27,37,42,43,49,62,66,67	
Establishing roles	Data custodian	The entity responsible for keeping data secure and confidential.	10,32,38,55	10,18,27,38,48,61	10,55
	Data steward	"The entity responsible for ensuring appropriate oversight of data use."[10]	5,10,46		10,46
	Data owner	The entity responsible for the raw data. In some instances, the data owner can also be the data custodian.	9,32,55	9,30	55
	Developers	The entity responsible for developing and monitoring the performance of the model.	9,32,38	19,33,36,37,49,62,65	
	Project management team	The team responsible for overseeing project planning and coordination activities.		50	51
Establishing and considering Health consumers	Community jury	A structure to support engagement and elicit health consumers views regarding health data collection, storage, sharing, and use.		18,27	
	Data subject	The "consumers from whom personal data are collected."[38]	38	38	

Blank cells: Mechanism not investigated for the specific techniques in the include studies.
FL Federated Learning, *ML* Machine Learning, *FDN* Federated Data Networks.

custodians, and stewards responsible for securing data[18] and overseeing its use[10]. These actors will be unlikely to leverage FL and associated techniques if they are not provided with clear directives and policies[10]. Regarding the governance of FL, the need to go beyond custodianship and consider the pertinent role of data stewards was also discussed[10]. Data stewards seek to maximise the benefits of data use while upholding data subjects' privacy and ensuring the data is not misused. They are also responsible for performing the analysis within their designated node. The data owner organisation or a trusted intermediary can perform this role. Developers are responsible for developing the predictive models[33] and, in some instances, software-assisted medical devices[37]. They are also responsible for the ongoing monitoring of performance[37]. Developers can be internal or external to the organisation where the data is collected. Regardless of where they are situated, developers need to maintain the privacy of the data they use[38]. Project management teams with project leads are recommended[50,51] to ensure the predictive models are developed effectively and efficiently.

During the development and use of ML, health consumers play a key role as evidenced by the relational mechanism of consumer involvement. However, many studies are silent on the role of health consumers, often referring to them as the data subject[38]. Others have demonstrated the utility of citizen juries in eliciting the views of health consumers[27]. The notion of consumer-driven data commons was also raised as an approach "to enable groups of consenting individuals to collaborate to assemble powerful, large-scale health data resources for use in scientific research, on terms that the group members themselves would set"[18].

Discussion

Synthesising this study's findings, Fig. 2 illustrates a conceptual framework of the procedural, relational, and structural mechanisms that should be considered when governing FL in healthcare. In illuminating these mechanisms, this study provides the first critical step to forming an evidence base to facilitate the effective governance of FL in healthcare. Our findings extend beyond previous literature on FL in healthcare in three ways:

First, in contrast to the predominant discourse in literature that views FL as a solution to the governance constraints surrounding the sharing of

sensitive health information[5], we identify that FL is accompanied by a raft of governance challenges. These governance challenges need to be addressed to realise the potential of FL whilst ensuring that it is appropriately used, safeguarding health consumers and health professionals. Failure to effectively govern FL risks deteriorating the consumers social licence regarding the secondary use of their health data[10] and compromises translating FL into clinical practice[52]. Healthcare organisations should take a balanced and measured approach to FL initiatives and recognise that leveraging FL may be accompanied by risks, but not leveraging FL will impede the necessary learnings to enhance the equitable and sustainable delivery of healthcare.

Second, we identified the procedural, relational, and structural mechanisms that healthcare organisations should enact when governing FL. The findings demonstrate that many of the governance mechanisms of FL are similar in intent to the governance mechanisms of traditional ML but may need to be enacted differently. For instance, due to the distributed nature of FL, with both the data and the learning residing locally within each healthcare organisation node, greater coordination efforts are required between the FL model provider and the nodes[10,32]. This also places a greater onus on capability development within the nodes to ensure the data stewards have the skills necessary to locally preprocess data and run the FL model according to a pre-defined and commonly agreed standard[32]. This also necessitates the new role of the 'data steward'[10], which is not apparent in traditional ML literature. We also identified that some ML governance mechanisms have not been considered in FL governance literature. Given that FL is a form of ML that centres on the secondary use of data and is governed by the same high-level principles (e.g. transparency, fairness, accountability)[10], it is likely that these mechanisms are relevant to FL but with nuances in their enactment. However, due to the papers investigating governance being largely conceptual or perspective papers, these nuances have yet to be examined and explored in real-world applications beyond limited exemplars. Future research, should seek to identify if these mechanisms are relevant and, if so, how they can be tailored for FL in healthcare.

Third, we provide a consolidated account that identifies and classifies governance mechanisms into theoretically meaningful and practically
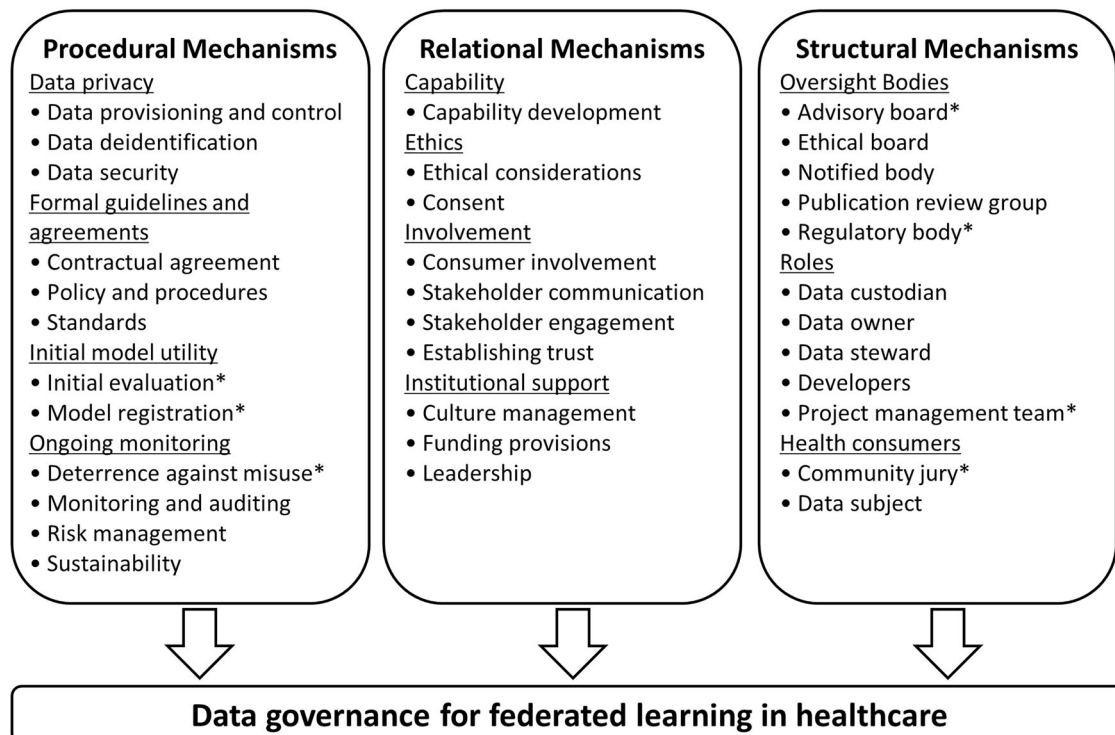

Procedural Mechanisms	**Relational Mechanisms**	**Structural Mechanisms**
<u>Data privacy</u> • Data provisioning and control • Data deidentification • Data security <u>Formal guidelines and agreements</u> • Contractual agreement • Policy and procedures • Standards <u>Initial model utility</u> • Initial evaluation* • Model registration* <u>Ongoing monitoring</u> • Deterrence against misuse* • Monitoring and auditing • Risk management • Sustainability	<u>Capability</u> • Capability development <u>Ethics</u> • Ethical considerations • Consent <u>Involvement</u> • Consumer involvement • Stakeholder communication • Stakeholder engagement • Establishing trust <u>Institutional support</u> • Culture management • Funding provisions • Leadership	<u>Oversight Bodies</u> • Advisory board* • Ethical board • Notified body • Publication review group • Regulatory body* <u>Roles</u> • Data custodian • Data owner • Data steward • Developers • Project management team* <u>Health consumers</u> • Community jury* • Data subject

Data governance for federated learning in healthcare

Fig. 2 | Procedural, relational, and structural mechanisms necessary to govern FL. Note: asterisk (*) denotes that the mechanism has yet to be explicitly reported in FL papers but are apparent in FDN/ML papers. Future research is necessary to identify their relevance and how they can be tailored to FL.

relevant categories (i.e., procedural, relational, and structural mechanisms). This overcomes the fragmented, piecemeal, and incomplete insights offered by existing literature on the governance of FL in healthcare. Only seven papers in our review examined FL governance in healthcare, and, as such, we complemented the insights with studies on the governance of the related techniques - ML and FDN. The limited research on FL data governance is also reflected in the immaturity of the literature on the governance of ML and FDN in healthcare. No single paper in our review examined all governance mechanisms. This immaturity also resembles research on healthcare governance in general, with a review finding that the literature offers few empirical insights on the efficacy of governance approaches[53]. In our review, we identified that studies often focused on elements surrounding governance rather than governance itself. For instance, studies often discuss regulations and standards imposed at the field level[49] rather than examine governance within and between organisations. Other studies highlight that health stakeholders' require digital and algorithmic literacy[40]. Yet, from a governance perspective, limited insights were provided regarding what this would entail. In addition, studies seldom empirically investigated the outcomes of the governance of FL and associated techniques, preferencing examination of the utility of the technical solution as opposed to the efficacy of the governance mechanisms. Future research should perform in-depth case studies focusing specifically on the governance of FL (and associated techniques), to identify specific mechanisms in use, how stakeholders enact them, and their outcomes. Such an approach is necessary to provide actionable insights to healthcare organisations considering or undergoing FL initiatives.

In terms of limitations, this review is scoped to focus on the governance of FL in healthcare, with insights from the governance of ML and FDN in healthcare supplementing the analysis. Focusing specifically on governance, a well-established term, the core search term was 'governance' without including related concepts. Whilst this is consistent with other reviews[21], the potential exists for articles discussing related concepts to be missed. In supplementing our findings on the governance of FL in healthcare with ML and FDN, additional techniques could be considered, including broader concepts of AI and distributed networks. However, with increasing breadth, the relevance of the insights to FL diminishes. The granularity of insights surrounding the governance mechanisms extracted in the Findings section was limited to the depth of insights provided in the included studies. Many of the included studies, while providing much-needed initial insights into the governance of FL or ML more broadly, only provide broad, high-level details surrounding the governance mechanisms. Future research should perform in-depth case studies on the governance of FL to offer more specificity on the underlying mechanisms. In addition, future research should also juxtapose this study's findings with an analysis of publicly available governance frameworks used in practice to synthesise procedural, relational, and structural mechanisms. National variations were not unpacked when developing a consolidated account of the mechanisms to consider when governing FL. Future research should seek to identify whether the governance complexities and mechanisms differ between different countries.

In summary, this research provides the critical first step to forming an evidence base to facilitate the effective governance of FL in healthcare. It provides healthcare organisations with key mechanisms to focus on when governing FL. In addition, this research provides the foundations necessary for cumulative research on facilitating the effective governance of FL in healthcare. Specifically, we call for further research to perform robust case studies to identify how to effectively enact each mechanism in practice. Such insights are critical as the healthcare industry progresses with the push to become a learning health system.

Scoping review methodology

This scoping review followed Arksey and O'Malley's five-stage approach[54]: 1) identifying the research question; 2) identifying relevant studies; 3) study selection; 4) charting the data; and 5) collating, summarising, and reporting the results. The subsections below detail how each stage was performed in this study. To ensure all relevant elements of the scoping review were adequately reported, as detailed in Supplementary Table 2, we adhered to the PRISMA-SCR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Extension for Scoping Reviews) Checklist.

Stage 1: Identifying research question

As elucidated in Arksey and O'Malley's five-stage approach, scoping reviews commence with clearly articulating the research question. As outlined in the Introduction, our research question was motivated by the necessity but lack of governance frameworks for FL in healthcare. Therefore, we were guided by the research question: What governance mechanisms can be used to underpin FL in healthcare?

Stage 2: Identifying relevant studies

A multipronged search with three independent search arms was performed (Table 4). The first search arm sought to provide insights into the governance mechanisms for FL in any context. The second search arm recognised that FL was a type of ML and sought to examine the governance of ML in healthcare. The third search arm examined the governance of federated data networks (FDN) in healthcare, recognising that FL is a variant of FDN (e.g. in FDN, data remains within the site, and metadata is accessible for others to query rather than 'learn'[55]). This multipronged search strategy was necessary because 1) there is limited research on the governance of FL; 2) FL is a form of ML and a variant of FDN, which involve the secondary use of data; and 3) the high-level principles that guide the governance of ML still remain relevant for the governance of FL. Failure to recognise the insights from the governance of ML and FDN could lead to a potentially overly simplistic and incomplete view of the factors to consider when governing FL. Across all search arms, and consistent with existing reviews[21], we leveraged the specific search term of 'governance' as it is a well-established concept in its own right. As discussed in Stage 4 and Stage 5 of the method, the governance mechanisms extracted were linked to the specific technique (e.g. FL, FDN, ML).

Multiple databases spanning multiple disciplines were searched, including healthcare, health informatics, computer science, and information systems. The databases searched were PubMed, EBSCOhost (MEDLINE and PsycINFO), ScienceDirect, Emerald Insight, Scopus, PubsOnLine, IEEE Xplore Digital Library, and the ACM Digital Library. The search was conducted in August 2023 and identified 672 articles.

Stage 3: Study selection

Two coders (RE, IC) independently applied inclusion and exclusion criteria to determine the articles' relevance (Table 5). During the abstract and full-text review, coder corroboration sessions were held until consensus was reached[56].

Table 4 | Search strategy

Search Arm	Purpose[a]	Search String[b]
1	Governing FL	(TITLE-ABS-KEY (governance) AND TITLE-ABS-KEY ("federated learning" OR "federated * learning"))
2	Governing ML in healthcare	(TITLE-ABS-KEY (governance) AND TITLE-ABS-KEY ("federated learning" OR "machine learning" OR "federated * learning") AND TITLE-ABS-KEY (healthcare OR health OR health* OR hospital* OR clinic* OR medical*))
3	Governing FDN in healthcare	(TITLE-ABS-KEY (governance) AND TITLE-ABS-KEY ("federated data" OR "federated database" OR "federated network") AND TITLE-ABS-KEY (healthcare OR health OR health* OR hospital* OR clinic* OR medical*))

[a]*FL* Federated Learning, *ML* Machine Learning, *FDN* Federated Data Networks.
[b]Supplementary Table 3 details the search string unique to each database.

Table 5 | Inclusion and exclusion criteria

Criteria	Exclusion Criteria	Inclusion Criteria
Relevance	• Provided a descriptive account of FL, ML, FDN without examining governance. • Developed or evaluated algorithms without examining governance. • Focused specifically on robotics. • Focused on digital technology broadly as opposed to FL or ML. • Examined specific contexts other than healthcare.	• Examined the governance (or related concepts) of FL in healthcare or general settings. • Examined the governance (or related concepts) of ML or FDN in healthcare.
Language	• Language other than English.	• English
Article types	• Conference proceeding abstracts, books, book chapters, theses, white papers, book reviews, presentations, tutorials, study protocols, preprints (without peer review). • Corrigendum, amendments, erratum, retracted articles. • Call for special issue.	• Peer reviewed journal and conference papers.

Table 6 | Classification framework informed by the conceptual framework for data governance[21]

Themes	Description
Procedural mechanisms	"Ensure that data is recorded accurately, held securely, used effectively, and shared appropriately" in accordance with policies, standards, and procedures[21].
Relational mechanisms	"Facilitate collaboration between team members" including communication, training, and shared values[21].
Structural mechanisms	"Determine reporting structures, governance bodies, and accountabilities" including decision making authorities and roles and responsibilities[21].

Stage 4: Charting the data

The articles were uploaded to NVivo (version 14; QSR International) and manually analysed using deductive and inductive analysis (RE, IC)[56]. During this stage, data was extracted regarding the region and industry in which the study was conducted, the technique examined (i.e. FL, ML, FDN), the nature of the study (i.e. conceptual or empirical), and the method employed.

The first author (RE) performed a backward and forward citation search[57] on the relevant FL governance articles, with a particular focus on identifying additional studies specifically related to FL governance. Backward and forward citation checking resulted in 241 and 40 additional papers, respectively, to consider for inclusion, of which 2 were included.

Stage 5: Collating, summarising, and reporting the results

For the deductive analysis, the Conceptual Framework for Data Governance[21] was used to capture the themes of procedural, relational, and structural mechanisms (Table 6). Within each theme, inductive coding was performed (RE, IC), where excerpts were extracted using verbatim codes and grouped based on similarities[58]. This resulted in 101 open codes (procedural mechanism open codes: $n = 38$; relational mechanisms open codes: $n = 27$, and structural mechanisms open codes: $n = 36$). Subsequently, constant comparisons of the open codes were performed[59]. This resulted in 34 mechanisms (procedural mechanisms: $n = 12$; relational mechanisms: $n = 10$; structural mechanisms: $n = 12$). Through additional constant comparison, the mechanisms were grouped into 11 subthemes (procedural subthemes: $n = 4$; relational subthemes: $n = 4$; structural subthemes: $n = 3$).

Each excerpt of text contained within the final 34 mechanisms was further analysed by extracting the technique (e.g. FL, ML, FDN) it was in reference to. This led to the development of Tables 1–3 in the Findings section, which identified what governance mechanisms have been mentioned about FL specifically and what additional governance mechanisms may need to be considered based on insights from ML and FDN. If a mechanism has not been explored for a specific technique, it does not mean that the mechanism is not important for consideration for that technique rather it simply represents that it has not received sufficient attention in the studies included in the review.

Steps were taken to ensure rigour and credibility. This included developing a coding rule book and performing coder corroboration sessions[56]. In addition, each coder reviewed the other's coding, and discrepancies were discussed until a consensus was reached[60].

Data availability

This scoping review uses peer-reviewed articles from a combination of open-access and subscription journals. Each article was analysed in NVivo (version 14; QSR International), and no additional datasets were generated. The search protocol and theme descriptions have been provided to facilitate replicability.

References

1. Friedman, C. et al. Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *J. Am. Med. Inform. Assoc.* **22**, 43–50 (2015).
2. Platt, J. E., Raj, M. & Wienroth, M. An analysis of the learning health system in its first decade in practice: scoping review. *J. Med. Internet Res.* **22**, e17026 (2020).
3. Syed, R. et al. Digital health data quality issues: systematic review. *J. Med. Internet Res.* **25**, e42615 (2023).
4. Austin, J. A. et al. Decades in the Making: The Evolution of Digital Health Research Infrastructure Through Synthetic Data, Common Data Models, and Federated Learning. *J. Med. Internet Res.* **26**, e58637 (2024).
5. Rieke, N. et al. The future of digital health with federated learning. *npj Digital Med.* **3**, 119 (2020).
6. Peng, L. et al. An in-depth evaluation of federated learning on biomedical natural language processing for information extraction. *npj Digital Med.* **7**, 127 (2024).
7. Abramson, W., Jones, H. A., Papadopoulos, P., Pitropakis, N. & Buchanan, W. J. A distributed trust framework for privacy-preserving machine learning. In *Trust, Privacy and Security in Digital Business: 17th International Conference, TrustBus*. 205-220 (2020).
8. Li, S. et al. Federated and distributed learning applications for electronic health records and structured medical data: a scoping review. *J. Am. Med. Inform. Assoc.* **30**, 2041–2049 (2023).
9. Li, Z., Mao, F. & Wu, C. Can we share models if sharing data is not an option?. *Patterns* **3**, 100603 (2022).
10. Suver, C., Thorogood, A., Doerr, M., Wilbanks, J. & Knoppers, B. Bringing code to data: Do not forget governance. *J. Med. Internet Res.* **22**, 11 (2020).

11. Xu, J. et al. Federated learning for healthcare informatics. *J. Healthc. Inform. Res.* **5**, 1–19 (2021).

12. Abbas, S. R., Abbas, Z., Zahir, A. & Lee, S. W. Federated Learning in Smart Healthcare: A Comprehensive Review on Privacy, Security, and Predictive Analytics with IoT Integration. In *Healthcare*. 2587 (2024).

13. Brisimi, T. S. et al. Federated learning of predictive models from federated electronic health records. *Int. J. Med. Inform.* **112**, 59–67 (2018).

14. Sheller, M. J., Reina, G. A., Edwards, B., Martin, J. & Bakas, S. Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In *International MICCAI Brainlesion Workshop*. 92-104 (Springer, 2019).

15. Roy, A. G., Siddiqui, S., Pölsterl, S., Navab, N. & Wachinger, C. Braintorrent: A peer-to-peer environment for decentralized federated learning. *arXiv Prepr. arXiv* **1905**, 06731 (2019).

16. Sharafaddini, A. M., Esfahani, K. K. & Mansouri, N. Deep learning approaches to detect breast cancer: A comprehensive review. *Multimedia Tools and Applications*, 1-112 (2024).

17. Moshawrab, M., Adda, M., Bouzouane, A., Ibrahim, H. & Raad, A. Reviewing federated machine learning and its use in diseases prediction. *Sensors* **23**, 2112 (2023).

18. Price, W. N. II & Chen, I. G. Privacy in the age of medical big data. *Nat. Med.* **25**, 37–43 (2019).

19. Solomonides, A. E. et al. Defining AMIA's artificial intelligence principles. *J. Am. Med. Inform. Assoc* **29**, 585–591 (2022).

20. Thoral, P. J. et al. Sharing ICU patient data responsibly under the society of critical care medicine/European society of intensive care medicine joint data science collaboration: The Amsterdam university medical centers database (AmsterdamUMCdb) example. *Crit. Care Med.* **49**, e563 (2021).

21. Abraham, R., Schneider, J. & Brocke, J. V. Data governance: A conceptual framework, structured review, and research agenda. *Int. J. Inf. Manag.* **49**, 424–438 (2019).

22. Winter, J. S. & Davidson, E. Governance of artificial intelligence and personal health information. *Digital Policy*. *Regul. Gov.* **3**, 280–290 (2018).

23. Wen, J. et al. A survey on federated learning: challenges and applications. *Int. J. Mach. Learn. Cybern.* **14**, 513–535 (2023).

24. Singh, P., Singh, M. K., Singh, R. & Singh, N. in *Federated Learning for IoT Applications* 199-214 (Springer, 2022).

25. AbdulRahman, S. et al. A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. *IEEE Internet Things J.* **8**, 5476–5497 (2020).

26. Papadopoulos, P., Abramson, W., Hall, A. J., Pitropakis, N. & Buchanan, W. J. Privacy and trust redefined in federated machine learning. *Mach. Learn. Knowl. Extraction* **3**, 333–356 (2021).

27. Degeling, C. et al. Community perspectives on the benefits and risks of technologically enhanced communicable disease surveillance systems: a report on four community juries. *BMC Med. Ethics* **21**, 1–14 (2020).

28. Ciampi, M., Sicuranza, M. & Silvestri, S. A privacy-preserving and standard-based architecture for secondary use of clinical data. *Information* **13**, 1–16 (2022).

29. Kellmeyer, P. Big brain data: On the responsible use of brain data from clinical and consumer-directed neurotechnological devices. *Neuroethics* **14**, 83–98 (2018).

30. Stevens, L. M. et al. American Heart Association precision medicine platform addresses challenges in data sharing. *Circulation: Cardiovascular Qual. Outcomes* **14**, 4 (2021).

31. Bardenheuer, K., Van Speybroeck, M., Hague, C., Nikai, E. & Price, M. Haematology outcomes network in europe (HONEUR)—A collaborative, interdisciplinary platform to harness the potential of real-world data in hematology. *Eur. J. Haematol.* **109**, 138–145 (2022).

32. Pati, S. et al. Federated learning enables big data for rare cancer boundary detection. *Nat. Commun.* **13**, 7346 (2022).

33. Bedoya, A. D. et al. A framework for the oversight and local deployment of safe and high-qualtiy prediction model. *J. Am. Med. Inform. Assoc.* **29**, 1631–1636 (2022).

34. Svensson, A. M. & Jotterand, F. Doctor ex machina: A critical assessment of the use of artificial intelligence in healthcare. *J. Med. Philos.* **47**, 155–178 (2022).

35. Wiljer, D. & Hakim, Z. Developing an artificial intelligence–enabled health care practice: Rewiring health care professions for better care. *J. Med. Imaging Radiat. Sci.* **50**, S8–S14 (2019).

36. Demir, E. Big Biological Data: Need for Reorientation of the Governance Framework. In *IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology*. 1-7 (2022).

37. Ho, C. W.-L. & Caals, K. A call for an ethics and governance action plan to harness the power of artificial intelligence and digitalization in Nephrology. *Semin. Nephrol.* **41**, 282–293 (2021).

38. Li, S.-C., Chen, Y.-W. & Huang, Y. Examining compliance with personal data protection regulations in interorganizational data analysis. *Sustainability* **13**, 11459 (2021).

39. Mello, M. M. & Wang, C. J. Ethics and governance for digital disease surveillance. *Science* **368**, 951–954 (2020).

40. Currie, G. & Hawk, K. E. Ethical and legal challenges of artificial intelligence in nuclear medicine. In *Seminars in Nuclear Medicine*. 120-125 (Elsevier, 2021).

41. Hine, C., Nilforooshan, R. & Barnaghi, P. Ethical considerations in design and implementation of home-based smart care for dementia. *Nurs. Ethics* **29**, 1035–1046 (2022).

42. Keen, J. et al. Machine learning, materiality and governance: a health and social care case study. *Inf. Polity* **26**, 57–69 (2021).

43. Carter, D. J. et al. Personal data for public benefit: The regulatory determinants of social licence for technologically enhanced antimicrobial resistance surveillance. *J. Law Med.* **30**, 179–190 (2023).

44. Keen, J. et al. Public Services, Personal Data and Machine Learning: Prospects for Infrastructures and Ecosystems. In *19th European Conference on Digital Government*. 51-54 (2019).

45. Yousefi, Y. Data Sharing as a Debiasing Measure for AI Systems in Healthcare: New Legal Basis. In *Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance*. 50-58 (2022).

46. Hallock, H. et al. Federated networks for distributed analysis of health data. *Front. Public Health* **9**, 712569 (2021).

47. National Center for Advancing Translational Sciences. (ed National Institutes of Health) (n.d.).

48. Nellåker, C. et al. Enabling global clinical collaborations on identifiable patient data: The Minerva Initiative. *Front. Genet.* **10**, 611 (2019).

49. Gilbert, S. et al. Learning from experience and finding the right balance in the governance of artificial intelligence and digital health technologies. *J. Med. Internet Res.* **25**, e43682 (2023).

50. Watson, J. et al. Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: What can we learn from US academic medical centers?. *JAMIA Open* **3**, 167–172 (2020).

51. Visweswaran, S. et al. Accrual to clinical trials (ACT): A clinical and translational science award consortium network. *JAMIA Open* **1**, 147–152 (2018).

52. Teo, Z. L. et al. Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture. *Cell Rep Med* **5**, 101419 (2024).

53. Pyone, T., Smith, H. & van den Broek, N. Frameworks to assess health systems governance: A systematic review. *Health Policy Plan.* **32**, 710–722 (2017).

54. Arksey, H. & O'malley, L. Scoping studies: Towards a methodological framework. *Int. J. Soc. Res. Methodol.* **8**, 19–32 (2005).

55. Alvarellos, M. et al. Democratizing clinical-genomic data: How federated platforms can promote benefits sharing in genomics. *Front. Genet.* **13**, 1045450 (2023).

56. Saldaña, J. *The coding manual for qualitative researchers*. (SAGE publishing, 2021).

57. Webster, J. & Watson, R. T. Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii (2002).

58. Dubois, A. & Gadde, L.-E. "Systematic combining"—A decade later. *J. Bus. Res.* **67**, 1277–1284 (2014).

59. Gioia, D. A., Corley, K. G. & Hamilton, A. L. Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organ Res Methods* **16**, 15–31 (2013).

60. Bandara, W. & Syed, R. The role of a protocol in a systematic literature review. *Journal of Decision Systems*, 1-18 (2023).

61. Schneider, G. Health data pools under European policy and data protection law: Research as a new efficiency defence?. *J. Intellect. Prop., Inf. Technol., Electron. Commer. Law* **11**, 49–67 (2020).

62. Li, P., Williams, R., Gilbert, S. & Anderson, S. Regulating AI/ML-enabled Medical Devices in the UK. In *Proceedings of the First International Symposium on Trustworthy Autonomous Systems*. 1-10 (2023).

63. Macrae, C. Governing the safety of artificial intelligence in healthcare. *BMJ Qual. Saf.* **28**, 495–498 (2019).

64. Agbese, M. et al. Governance of ethical and trustworthy al systems: Research gaps in the eccola method. In *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*. 224-229 (IEEE, 2021).

65. Ho, C. W.-L. Deepening the normative evaluation of machine learning healthcare application by complementing ethical considerations with regulatory governance. *Am. J. Bioeth.* **20**, 43–45 (2020).

66. Wellnhofer, E. Real-world and regulatory perspectives of artificial intelligence in cardiovascular imaging. *Front. Cardiovascular Med.* **9**, 890809 (2022).

67. Ho, C. W.-L. & Malpani, R. Scaling up the research ethics framework for healthcare machine learning as global health ethics and governance. *Am. J. Bioeth.* **22**, 36–38 (2022).

Acknowledgements

Author contributions

RE and CS devised the project. RE and IC led the search and the analysis. CB, SB, LC, SdJ, YG, A-DG, ML, PM, AN, SMP, MS, CS provided comprehensive feedback and recommendations of the search strategy. RE led the writing of the manuscript with comprehensive edits, feedback and recommendations provided by all other authors (IC, CB, SB, LC, SdJ, YG, A-DG, ML, PM, AN, SMP, MS, CS). All authors contributed to the article and approved the submitted version.

Competing interests

Anthony Nguyen is an Associate Editor of npj Digital Medicine. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-025-01836-3.

Correspondence and requests for materials should be addressed to Rebekah Eden.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.