

Ibrahim Dhouafli ***is a doctoral researcher at Brunel Design School, Brunel University of London, where his thesis examines civil liability, AI accountability and regulatory governance in AI-mediated clinical decision-making systems, drawing comparative insights from the United Kingdom and Tunisia. His work develops the concept of Digital Injury© as a category of compensable harm and proposes a responsible design framework linking human oversight duties to patient compensation. He is supervised by Dr Federico Colecchia, Dr Zoi Krokida and Dr Gabriella Spinelli . His wider research interests span health law, technology regulation and socio-legal design methods. Orchid id: 0009-0007-7074-8603 Emial: Ibrahim.dhouafli@brunel.ac.uk.***



Dr Federico Colecchia ***(PhD in Physics, 2001) is Senior Lecturer in Creative Electronics and Programming at Brunel University of London. He is a specialist in emerging technologies and sustainable development with 20+ years of research experience across academia and the private sector. His current work has a focus on human-centred design of innovations to promote individuals' health and wellbeing relying on applications of artificial intelligence and immersive technologies. Orchid ID: 0000-0001-7447-711. Email: federico.colecchia@brunel.ac.uk***

ISSN: 2582 – 8304

Towards Responsible Design for Equitable AI-Mediated Clinical Decision-Making Systems: A Scoping Review of Health Equity and Underserved Populations

Ibrahim Dhouafli[1],

Federico Colecchia[2]

Abstract

Artificial intelligence is becoming a core feature of Clinical Decision-Making Systems for diagnosis, risk prediction and treatment planning. However, the benefits are not equally distributed and underserved communities remain at risk from algorithmic bias, accessibility barriers and digital exclusion.

This scoping review synthesises evidence on health equity within AI-mediated Clinical Decision-Making Systems from a Responsible Design perspective, with emphasis on racialised communities, disabled patients and digitally excluded populations.

The review followed the Arksey and O'Malley framework (2005) and reported in accordance with the PRISMA-ScR. Scopus was searched iteratively. Overall eligibility covered English-language publications indexed in Scopus and published between 2016 and 2026. 2,082 records were found. After deduplication, relevance screening and structured evidence mapping across seven predefined conceptual domains, 80 publications were selected and analysed using thematic synthesis.

Evidence was dominated by health equity (74/80) and algorithmic bias (55/80). Inequality occurred throughout the entire AI lifecycle, from data collection and model development to implementation, deployment and monitoring, not only from algorithmic performance. Racialised communities received the most attention, while disability (8/80), digital exclusion (23/80) and legal liability (6/80) were underrepresented. Responsible Design was addressed in 79/80 studies, but few translated it into operational design or governance practices. Governance remained fragmented across disciplines and jurisdictions, post-deployment equity assessment was scarce, and underserved communities rarely had a role in system design or governance.

AI-mediated clinical decision-making may reproduce and amplify existing inequalities when deployed within health systems characterised by structural weaknesses. These risks are significant for underserved communities where limitations in data representation, digital access, inclusive participation and infrastructure remain unresolved. Equitable implementation requires these systemic conditions to be addressed alongside AI design, deployment and governance. AI should therefore be understood not as a solution to existing inequities, but as a technology whose effects depend on inclusiveness and quality of systems.

Keywords: ***artificial intelligence; Clinical Decision-Making Systems; health equity; Responsible Design; underserved populations; algorithmic bias; governance; Liability; scoping review***

Introduction

The rapid integration of AI into Clinical Decision-Making Systems (CDMS) is transforming diagnosis, risk prediction and treatment planning (Bashir et al., 2026; Wen et al., 2024). As these systems become routine in clinical practice, they have evolved from technological innovations into socio-technical systems that influence how care is delivered, accessed and distributed across populations (Belmejdoub et al., 2026; Cappello et al., 2026).

Health equity is shaped by social and socioeconomic factors like income, education, discrimination and healthcare access. These factors influence both real-world health outcomes and the clinical data used to train AI systems (WHO, 2025). Consequently, AI trained on biased or unrepresentative data may treat patients "equally" while producing inequitable outcomes. This underscores the difference between equality (treating everyone the same) and equity (allocating resources based on need), where health inequities represent avoidable and unjust differences (Whitehead, 1992; Whitehead & Dahlgren, 2006). Here, health equity is defined as the absence of systematic, remediable differences in health outcomes or resource access between socially advantaged and disadvantaged groups (Braveman & Gruskin, 2003).

Because AI-mediated clinical decision-making systems (CDMS) learn from historical data within unequal healthcare environments, underserved communities need specific safeguards (Hoffman, 2021; Tsai et al., 2022). Algorithmic bias, underrepresentation, and digital exclusion create inequitable outcomes that disproportionately impact vulnerable populations (Budhu et al., 2025; Hilling et al., 2025). Specifically, racialised groups face lower diagnostic precision from

imbalanced datasets (Chin et al., 2023), disabled patients encounter accessibility barriers (Olawade et al., 2025; Ricchezze et al., 2025), and digitally excluded populations experience reduced access to AI healthcare (Osonuga et al., 2025).

Research in technical fields has addressed bias mitigation, algorithmic fairness, and explainable, trustworthy AI (Al-Duais et al., 2026; Lan et al., 2025). Design research promotes human-centred, inclusive, and participatory methods, involving clinicians, patients, and community representatives in technology development (Bokolo et al., 2026; Clifford et al., 2022). Legal, ethical, and policy researchers have established governance frameworks for accountability, transparency, and regulatory oversight (Cary et al., 2025; Lau, 2025).

Fragmented evidence and isolated designs create a critical gap in equitable AI systems for underserved communities. Existing studies usually evaluate demographic groups separately, neglecting how interconnected factors impact health equity in AI-mediated CDMS (Belmejdoub et al., 2026; Ingraham et al., 2023). While some reviews address specific healthcare AI aspects (Adegunle et al., 2026; Murphy et al., 2025), none integrate these via a Responsible Design lens for underserved groups.

Unlike isolated approaches to fairness, accessibility, transparency, or governance, Responsible Design provides an integrated solution. It advances equitable technology by combining ethical, organizational, technical, and societal factors throughout system development (Arockiasamy, 2025; Bagali et al., 2026). This framework promotes stakeholder engagement, inclusivity, accountability, and real-world system context during implementation, monitoring, and continuous improvement (Bokolo et al., 2026; Dinc and Ardic, 2026).

This scoping review synthesises the current evidence on health equity within AI-mediated CDMS from a Responsible Design perspective, addressing four questions: (1) How does the literature conceptualise health equity within AI-mediated CDMS? (2) What equity challenges are experienced by racialised communities, disabled patients and digitally excluded populations? (3) How are Responsible Design, governance and accountability currently incorporated into AI-mediated CDMS? (4) What evidence gaps and future research priorities emerge from the existing literature?

2. Methods

2.1 Study Design

A scoping review method was selected for synthesising literature on health equity in AI-mediated CDMS, emphasising underrepresented communities and Responsible Design, given the field's multidisciplinary nature. The review followed Arksey and O'Malley framework and was reported according to PRISMA-ScR. No protocol was prospectively registered.

Scopus was searched through 16 iterative searches combining AI-mediated CDMS terms with equity, harm, design, or governance concepts. Eligibility covered English publications (2016–2026), yielding 2,082 records (1,697 unique after removing 385 duplicates: 365 by DOI, 20 by title/year).

Titles and abstracts were evaluated against the predefined eligibility criteria. Eligible records were assessed across seven conceptual domains: AI-mediated CDMS, health equity/disparities, underserved communities, Responsible Design, governance/accountability, methodological relevance, and publication type. For the first six, predefined relevance terms identified in the title, author-supplied

keywords recorded in Scopus, and the abstract were weighted 3, 2, and 1 points, respectively, with each term counted once at its highest-value location. Domain scores were capped at 12, converted to a 0–4 scale, and multiplied by specific domain weights: 2.5 (CDMS), 3.0 (health equity), 3.0 (underserved), 2.0 (Responsible Design), 1.5 (governance), and 1.0 (methodology), totalling up to 52 points. The seventh domain, publication type, contributed a source-type score of 1.0 for journal articles and 0.5 for conference papers and book chapters. A post-2020 recency bonus of 0.5 points per year, capped at 3.0, was added. The resulting weighted relevance score had a maximum of 56 points, with 15.0 used as the minimum relevance threshold. Records meeting the threshold were subsequently ranked, and the 80 highest-ranked publications were retained.

Table 1: Evidence Mapping Framework

Conceptual Domain	*Purpose*
AI-mediated CDMS	***Relevance to clinical AI***
Health equity and healthcare disparities	***Focus on equity outcomes***
Underserved communities	***Inclusion of target populations***
Responsible Design	***Design principles and approaches***
Governance and accountability	***Oversight, regulation and responsibility***
Methodological relevance	***Contribution to the review objectives***
Publication type	***Contribution to relevance ranking based on publication type***

2.2 Eligibility Criteria

This review evaluated literature on AI-mediated CDMS and health equity. During screening, papers were eligible if they investigated CDMS design, implementation, governance, or evaluation while

addressing health equity, algorithmic fairness, healthcare disparities, underserved populations, or Responsible Design—particularly racialised, disabled, or digitally excluded groups. These screening criteria were not encoded in the search queries.

Two other parameters bounded the corpus. First, all records came from Scopus, a multidisciplinary database spanning healthcare, computer science, and social sciences. Second, the 2016–2026 publication window varied by search: only six of sixteen searches used the full range, ten started later, and one ended in 2025. Thus, early year coverage is uneven; themes like governance, liability, and automation bias had shorter windows than health equity, bias, and underserved populations.

Table 2. Eligibility criteria.

Criterion	*Inclusion*	*Exclusion*
Population	***Underserved communities and other vulnerable groups***	***Populations outside the review scope***
Context	***AI-mediated CDMS***	***Non-clinical AI or general AI applications without healthcare relevance***
Concept	***Health equity, algorithmic fairness, Responsible Design***	***Purely technical algorithm development without clinical or equity implications***
Study type	***Peer-reviewed research articles, systematic and scoping reviews, conference papers, book chapters, and other scholarly publications***	***Editorials, commentaries, opinion pieces, and non-scholarly or non-peer-reviewed material without sufficient relevance to the review questions***

Criterion	*Inclusion*	*Exclusion*
	relevant to the review questions	
Language	***English***	***Non-English***
Timeframe	**Published 2016–2026**	**Published before 2016**
Information source	**Indexed in Scopus**	**Not indexed in Scopus**
These criteria defined the final corpus eligibility rather than identical parameters for every retrieval search.		

2.3 Study Selection

Study selection followed PRISMA-ScR framework. After screening titles and abstracts against eligibility criteria, remaining records underwent full-text assessment for relevance to AI-mediated CDMS, health equity, underserved communities, and Responsible Design. Eligible records were scored across seven conceptual domains using weighted criteria based on title, abstract, and author-supplied keywords recorded in Scopus. Records meeting a 15.0-point minimum threshold were ranked by score. The top 80 records were selected and grouped into three priority tiers: Core (27), Important (27), and Relevant (26). The final ranked list was verified by the author.

2.4 Data Charting

Data were extracted using a charting format developed for this review. For each publication, key information was collected on publication details, study design, healthcare context, AI application, target population and geographical setting, health equity issues, algorithmic bias, accessibility, disability, digital exclusion,

governance, liability and Responsible Design approaches, together with the main findings and research gaps identified.

2.5 Data Synthesis

Thematic synthesis used a hybrid deductive–inductive approach. The initial data-charting framework captured equity, algorithmic bias, disability, digital exclusion, Responsible Design, governance, liability, care quality, and healthcare access. Extracted evidence was repeatedly examined to identify recurring patterns, differences and relationships across studies, allowing codes to be refined and grouped into broader themes. The final nine themes were retained as they consistently represented the principal issues identified across the evidence and provided a coherent basis for synthesis. For each theme, the evidence was synthesised to identify barriers to equitable AI implementation, current design approaches and improvement opportunities.

3. Results

3.1 Characteristics of Included Studies

The search and screening process revealed 80 publications that met the inclusion criteria from an initial 2,082 records, illustrated in the PRISMA-ScR flow diagram (Figure 1).

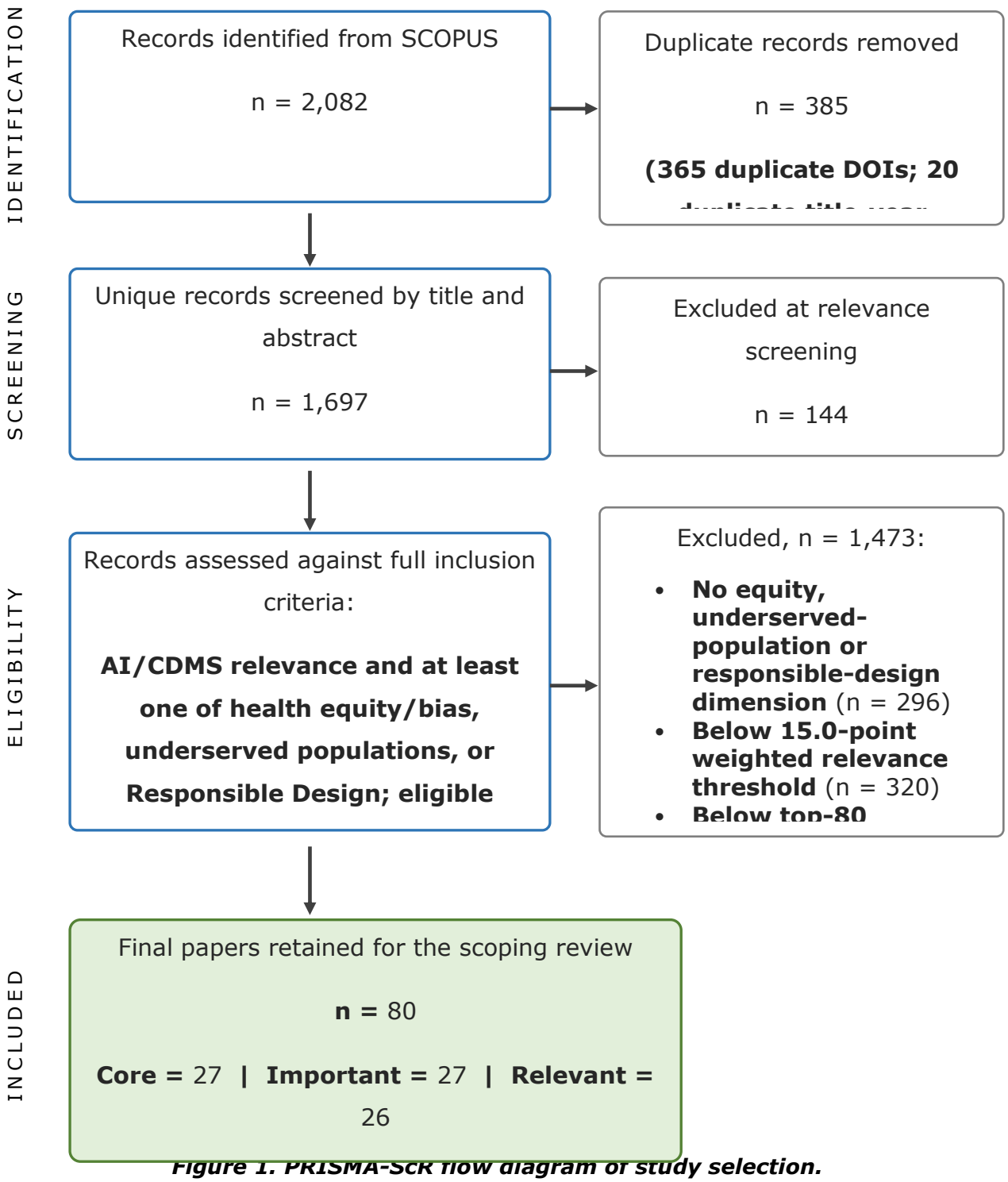


Figure 1. PRISMA-ScR flow diagram of study selection.

Note: The 15.0 threshold functioned as an initial relevance floor; it was not the criterion determining the final sample size. Records meeting the threshold were subsequently ranked by their total weighted relevance score, and the 80 highest-ranked records were retained.

Most included studies were published between 2022 and 2026 (Figure 2), reflecting the growing use of AI in healthcare and increasing attention to how algorithmic systems may reinforce existing inequalities when equity is not embedded in AI design, implementation and governance (Murphy et al., 2025; Zarnegar, 2026).

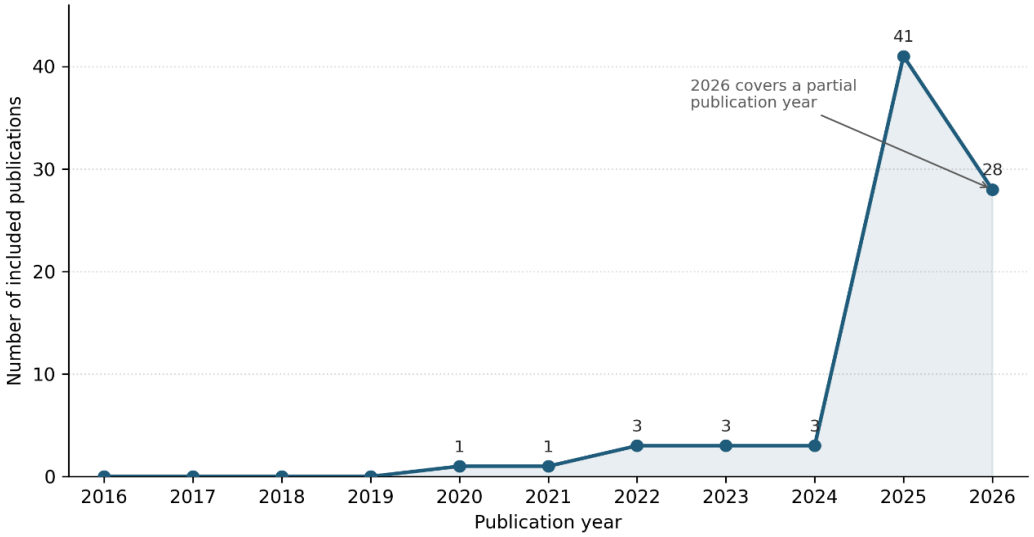


Figure 2. Included publications per year, 2016–2026 (N = 80).

The evidence base was methodologically diverse. Empirical studies focused on AI in diagnostic decision-making, clinical risk prediction, medical imaging, electronic health record analytics and patient triage (Al-Duais et al., 2026; Kukkar et al., 2026). Reviews examined algorithmic fairness, ethical governance and disparities (Adegunle et al., 2026; Murphy et al., 2025). Conceptual papers explored Responsible Design, inclusive governance (Bokolo et al., 2026; Ekuta, 2025).

Table 3. Characteristics of the included studies (N = 80).

Characteristic	*n*	*%*
Study design		

Narrative and literature reviews	***24***	***30.0***
Conceptual, perspective and policy papers	***18***	***22.5***
Technical/empirical model studies	***17***	***21.3***
Systematic and scoping reviews	***10***	***12.5***
Qualitative and mixed-methods studies	***10***	***12.5***
Other	***1***	***1.3***
Publication year		
2016–2022	***5***	***6.3***
2023–2024	***6***	***7.5***
2025	***41***	***51.3***
2026	***28***	***35.0***
Geographical context		
USA	***18***	***22.5***
Europe	***14***	***17.5***
of which United Kingdom	***(5)***	***(6.3)***
India / India-focused	***11***	***13.8***
Other countries	***20***	***25.0***
Not reported (global-scope reviews and conceptual papers)	***17***	***21.3***
AI application		
LLM, NLP and generative AI	***22***	***27.5***
Predictive and risk-prediction models	***17***	***21.3***
Diagnostic and screening AI	***17***	***21.3***
CDMS	***13***	***16.3***
Telemedicine, mobile and wearable AI	***11***	***13.8***

Most included studies focused on high-income-country settings, particularly the USA and Europe, including the United Kingdom (Adegunle et al., 2026; Hilling et al., 2025). Research from India

outlines increased interest in AI alongside persistent inequality in healthcare access, digital infrastructure and socio-economic status (Baruah and Majumdar, 2026; Roy et al., 2025).

Across the literature, AI-mediated CDMS emerged as socio-technical systems in which health equity is shaped by systems-level barriers including data quality, technical and algorithmic performance, accessibility, clinician oversight, digital literacy and institutional governance (Belmejdoub et al., 2026; Cappello et al., 2026).

Health equity, algorithmic bias and fairness dominated the corpus, while disability, digital exclusion, access to support and legal accountability were underrepresented (Hoffman, 2021; Olawade et al., 2025). Similarly, few studies examined lifecycle Responsible Design, favouring technical performance over its impact on people and society (Bagali et al., 2026; Ingraham et al., 2023).

3.2 Health Equity Challenges in AI-mediated CDMS

Health equity emerged as a central topic in the literature. AI-mediated CDMS can potentially improve healthcare but also contribute to inequalities through algorithmic design, healthcare systems, and broader social determinants of health (Iloanusi and Chun, 2024; Murphy et al., 2025; Lau, 2025).

Table 4. Thematic synthesis: findings and evidence gaps across the 80 included studies.

Theme	*Papers (n)*	*Main findings*	*Remaining gaps*
Responsible Design	***79***	***Participatory/co-design, explainable AI, accountability; principles shifting to operational tooling.***	***Limited framework replication; no reporting standard.***

Health equity	***74***	***Data representation and service access dominate outcomes.***	***Often aspirational; rarely measured; limited post-deployment evidence.***
Algorithmic bias	***55***	***Mainly linked to unrepresentative training data.***	***Mitigation rarely validated; intersectional bias overlooked.***
Quality of care	***55***	***Improved diagnosis, personalisation and care continuity***	***Few subgroup or longitudinal evaluations.***
Governance	***50***	***EU AI Act, US and WHO guidance dominate.***	***Largely conceptual; equity auditing remains aspirational.***
Access to support	***44***	***Telemedicine, mobile health and conversational AI expand access.***	***Outcomes rarely evaluated.***
Digital exclusion	***23***	***Literacy, connectivity, device access as structural preconditions of inequity***	***Mainly descriptive; affordability underexplored.***
Disability	***8***	***Assistive technology, rehabilitation, inclusive design.***	***Limited intersectional analysis with race, income and geography; little outcome data***
Liability	***6***	***Least developed; clinician liability concerns.***	***No adopted liability framework for clinical AI harm.***

Algorithmic bias was the primary source of inequity, driven by biased training data, underrepresentation, and unequal model performance across demographic groups (Al-Duais et al., 2026; Bai et al., 2026). These limitations are rooted in historical healthcare disparities, diagnosis and treatment. AI systems can inherit, reproduce and amplify existing inequalities (Hoffman, 2021; Tsai et al., 2022). Fairness therefore extends beyond algorithmic performance, as

higher accuracy cannot by itself address inequalities rooted in institutional structures (Husebo et al., 2025; Sarkar et al., 2025). Because inequities may emerge across the AI lifecycle, continuous governance is required rather than isolated technical interventions (Hasanzadeh et al., 2025; Chandra et al., 2025; Lan et al., 2025). Finally, although AI can improve diagnostics and decision support, its benefits are not necessarily distributed equitably (Bray et al., 2025; Wong et al., 2025).

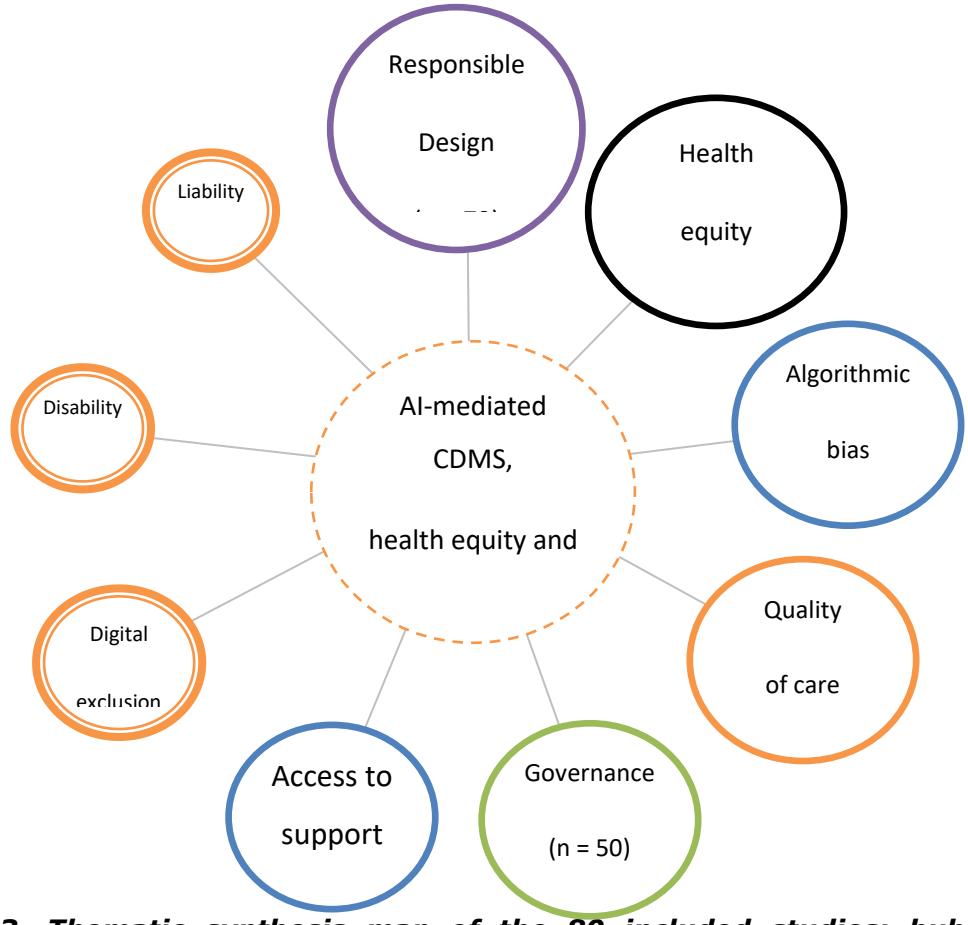


Figure 3. Thematic synthesis map of the 80 included studies; bubble size is proportional to the number of studies coded to each theme.

3.3 Experiences of Underserved Communities

The literature showed that the impact of AI-mediated CDMS is unequal, disproportionately affecting communities facing structural healthcare barriers (Budhu et al., 2025; Olawade et al., 2025). The

descriptor "Underserved" is contextual rather than universal, reflecting local histories and health systems conditions: it may apply to minorities (Chin et al., 2023; Hoffman, 2021), rural and linguistic groups (Baruah and Majumdar, 2026; Roy et al., 2025; Dinc and Ardic, 2026; Mwogosi, 2025), migrants (Cheong and Kubota, 2026), sexual minorities (Racine, 2025) and other populations experiencing structural barriers to healthcare. These groups share data underrepresentation, limited digital access, and exclusion from system design (Bokolo et al., 2026; Nair et al., 2025). Understanding these experiences depends on cultural, linguistic, and systemic contexts rather than a single profile.

3.3.1 - Racialised Communities

Racial and ethnic inequities were the most discussed issues in the literature (Ekuta, 2025; Murphy et al., 2025). AI systems can reproduce biases present in initial clinical data, with underrepresented populations potentially experiencing poorer outcomes (Hasanzadeh et al., 2025). These disparities reflect not only algorithmic limitations but broader inequalities in healthcare access, delayed diagnosis and social disadvantage (Hilling et al., 2025; Hoffman, 2021). The literature widely recognises demographic representation as essential (Chin et al., 2023; Zarnegar, 2026) and highlights that equitable AI requires continuous monitoring, transparent reporting and meaningful engagement with affected communities (Budhu et al., 2025; Clifford et al., 2022).

3.3.2 - Disabled Patients

Disability received less attention than racial inequities within the final corpus despite its importance in Responsible Design research (Olawade et al., 2025; Ricchezze et al., 2025). Most current research

focused on digital health rather than AI-mediated CDMS (Liu et al., 2025; Thorat and Kalani, 2026), and AI systems often disadvantage users with physical, sensory or cognitive impairments (Murugesan and Viswanathan, 2024; Olawade et al., 2025). The literature supports more inclusive design, but empirical evidence remains limited (Ricchezze et al., 2025; Thorat and Kalani, 2026).

3.3.3 - Digitally Excluded Populations

Digital exclusion was also a major contributor to inequitable AI-mediated CDMS (Fanarioti and Karpouzis, 2025; Nair et al., 2025). The main causes are device access, language barriers, poor connectivity, poverty, and limited literacy (Cheong and Kubota, 2026; Mwogosi, 2025). As a result, digital reliance creates new barriers that reflect broader structural inequalities (Kannadasan, 2026; McGrath et al., 2025; Mendez-Lopez et al., 2026).

3.3.4 - Cross-cutting Patterns Across Underserved Communities

Across all three groups, inequitable AI outcomes resulted not only from algorithmic bias but from the interaction of technical design with broader social and healthcare inequalities (Belmejdoub et al., 2026; Iloanusi and Chun, 2024). Underrepresentation and lack of community participation in AI governance were common concerns (Chin et al., 2023; Clifford et al., 2022). Researchers must adopt socio-technical approaches to address structural inequalities across the AI lifecycle, with systematic bias identification and mitigation extending from model conception through deployment and post-deployment surveillance (Hasanzadeh et al., 2025; Cappello et al., 2026; Lan et al., 2025).

3.4 Responsible Design, Governance and Accountability

Through the reviewed literature, Responsible Design emerged as a promising perspective to address the ethical and societal implications of AI-mediated CDMS (Cappello et al., 2026; Ekuta, 2025) (Figure 4). Many studies stressed fairness, inclusivity, transparency and human-centred development (Arockiasamy, 2025; Fesabas and Shaker, 2025), but few translated these principles into practical approaches supporting equitable CDMS (Bagali et al., 2026; Cary et al., 2025).

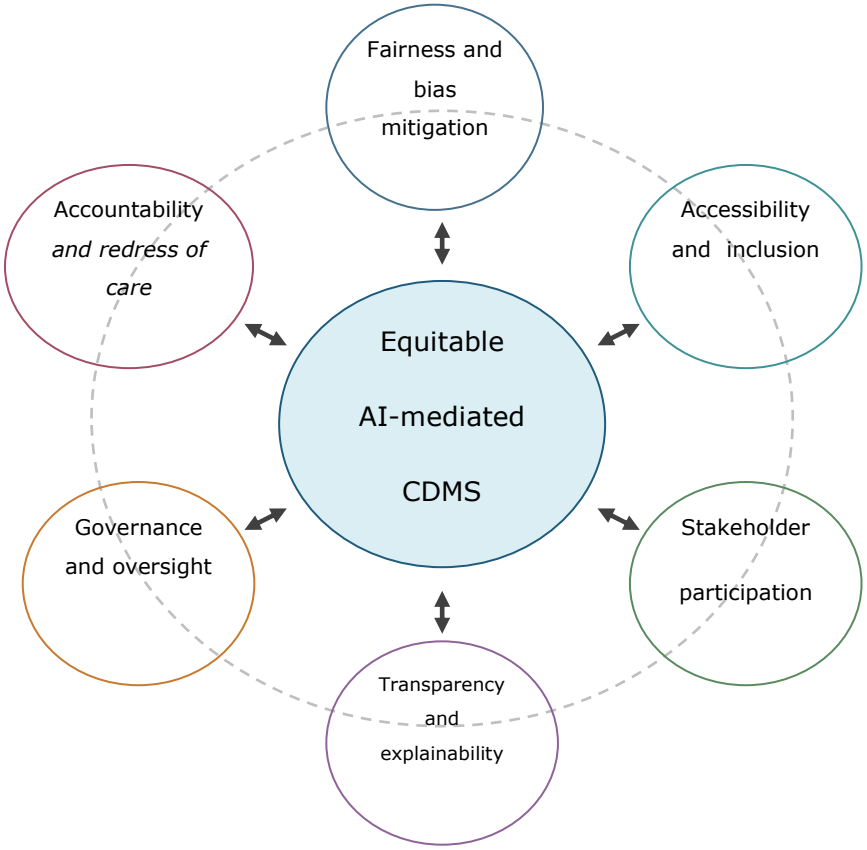


Figure 4. The Responsible Design perspective for equitable AI-mediated CDMS.

Responsible Design is described as a multidisciplinary approach integrating technical performance with broader ethical, social and organisational responsibilities, rather than a single methodology

(Arockiasamy, 2025; Radanliev, 2025). Equitable AI is not limited to algorithmic accuracy, it depends on understanding who benefits, who can be disadvantaged and how inequities can be addressed during development, deployment and evaluation (Bagali et al., 2026; Ekuta, 2025).

Participatory and human-centred approaches were widely recognised as core components of Responsible Design (Bokolo et al., 2026; Budhu et al., 2025). Several studies recommended clinicians, patients and community representation throughout AI-CDMS development to improve transparency, trust and inclusivity (Cheong and Kubota, 2026; Liu et al., 2025). However, this involvement was limited to system evaluation rather than the full AI lifecycle (Bokolo et al., 2026; Ingraham et al., 2023). Transparency and explainability also appeared as preconditions for trustworthy AI and professional accountability (Abu Assab et al., 2026; Sarkar et al., 2025).

The literature considers Responsible Design an organisational responsibility requiring constant monitoring of algorithm performance and unexpected consequences after deployment (Bagali et al., 2026; Wilkie et al., 2025). Formal governance is needed, with ethical guidelines, frameworks and policy initiatives to ensure fairness, transparency, human oversight and accountability. Within the deliberately targeted UK–India evidence, studies indicated differences in governance maturity but common gaps in ethical implementation (Baruah and Majumdar, 2026; Yu et al., 2026). Legal accountability is underrepresented in the final corpus, with few studies investigating responsibility or patient redress when AI leads to inequitable outcomes (Abu Assab et al., 2026; Savastano et al., 2025).

4. Discussion

This review synthesised 80 studies on health equity in AI-mediated CDMS through a Responsible Design perspective. Its main finding is that inequity is not solely a technical problem but a systemic one, produced across the AI lifecycle and amplified by fragmented governance.

4.1 Assessing the Gap: Why Advanced CDMS May Fail to Reduce Health Inequities?

The persistence of health inequities remains difficult to explain because research has not yet fully established where inequities arise across the clinical AI lifecycle or how different contributing factors interact. Most work has been oriented towards improving algorithmic accuracy rather than understanding how these systems work in real-clinical practice (Ingraham et al., 2023; Nouis et al., 2025).

Equity is an emergent property of the broader clinical ecosystem rather than the algorithm itself (Belmejdoub et al., 2026; Cappello et al., 2026). Because bias can arise throughout the AI lifecycle and poor monitoring harms performance, a model might pass a fairness audit but still cause inequitable outcomes (Hasanzadeh et al., 2025; Radanliev, 2025; Tsai et al., 2022). Ongoing model evaluation must look beyond metrics like classification accuracy to assess impact across diverse communities, especially marginalized groups (Chin et al., 2023; Ingraham et al., 2023). An intersectional approach helps clarify how systemic barriers like socio-economic status, ethnicity, and disability are interwoven (Lau, 2025; Putri, 2026). Most importantly, this evaluation must be continuous, spanning from model development to post-deployment surveillance (Al-Duais et al., 2026; Bagali et al., 2026).

4.2 Responsible Design as an Integrating Perspective for Equitable AI

A recurring pattern across the literature is a quiet emergent consensus. Studies on responsible AI, inclusive, human-centred, participatory design, and responsible innovation share one idea: equitable AI must actively protect the people impacted by technological decisions at all lifecycle stages (Arockiasamy, 2025; Clifford et al., 2022). This review interprets this consensus as a bridge to Responsible Design, not as one more methodology but as the perspective within which various approaches become synergistic: human-centred design is based on clinical and patient relevance (Giri et al., 2025; Gupta and Jaiswal, 2025); participatory and co-design involve local communities in designing systems rather than receiving them (Bai et al., 2026; Iloanusi and Chun, 2024); inclusive design extends usability beyond the average user (Bokolo et al., 2026; Thorat and Kalani, 2026). Responsible Design provides an integrating perspective that makes a place for fairness, accessibility, participation, transparency and accountability, the latter remaining comparatively underdeveloped in the literature (Figure 5).

The corpus demonstrates how far this integration is from being implemented. Some studies reduce Responsible Design to ethical principles like fairness and transparency (Fesabas and Shaker, 2025; Radanliev, 2025), others regard it as a participatory process with clinicians, patients or communities (Clifford et al., 2022; Liu et al., 2025) and only a few extend it to governance, monitoring and organisational learning after deployment (Bagali et al., 2026; Cary et al., 2025), highlighting a critical deficit in institutional infrastructure.

This is why Responsible Design addresses the gap discussed in Section 4.1. If inequity arises across a socio-technical lifecycle, the response should be distributed across it: inclusive datasets and explainability relevant to clinical reasoning at development (Marvin et al., 2023; Nouis et al., 2025), participatory processes at design, and mechanisms to monitor performance and adapt governance after deployment (Cappello et al., 2026; Ingraham et al., 2023). Responsible Design is thus an ongoing process, not a one-time checklist.

A structured approach to move from principle to practice is still lacking. Design Research Methodology provides this by cycling through research clarification, descriptive, prescriptive, and further descriptive studies (Blessing and Chakrabarti, 2009). This review aids the initial descriptive stage by synthesising how equity, harm, and governance interact across the AI lifecycle and lays the groundwork for a prescriptive stage in which Responsible Design artefacts can be developed and tested.

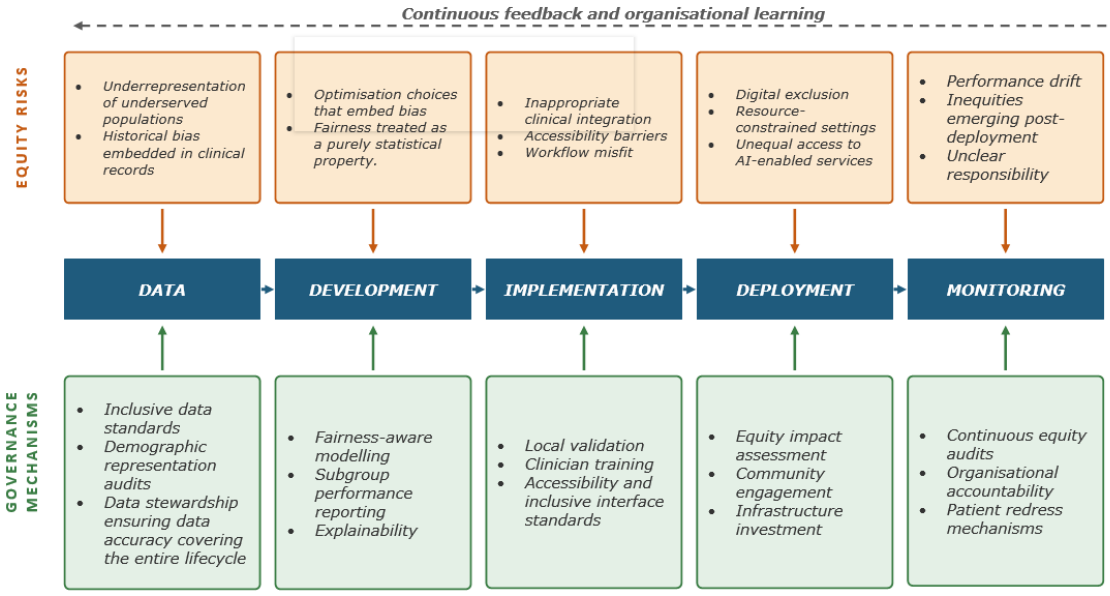


Figure 5. Responsible Design across the AI lifecycle: equity risks and governance mechanisms.

4.3 Governance and Accountability Beyond Technical Fairness

Governance was discussed in most studies (50/80) where the evidence remains fragmented. Ethical frameworks focused on fairness, transparency and autonomy (Fesabas & Shaker, 2025; Ricchezze et al., 2025). Regulatory guidance emphasised compliance and safety (Adegunle et al., 2026; Lacmanovic & Skare, 2025). Technical research focused on bias, explainability and model performance (Lan et al., 2025; Sarkar et al., 2025). Healthcare governance concentrated on clinical quality and patient safety (Cary et al., 2025; Nouis et al., 2025). Few studies integrated these perspectives into a single framework over the whole AI lifecycle (Bagali et al., 2026). This fragmentation is exacerbated by jurisdictional differences, as developers navigate frameworks such as the GDPR and the EU AI Act. Neither framework was specifically designed as a comprehensive framework for addressing health equity in clinical AI (Adegunle et al., 2026; Bagali et al., 2026).

A focused comparison between India and the UK, both deliberately targeted in the search strategy, is instructive. UK studies show strict rules exist, but new approaches are needed to ensure accountability and public inclusion in decision-making (Nouis et al., 2025; Racine, 2025). India studies highlight a struggle to build governance capacity as adoption speeds up alongside disparities (Baruah and Majumdar, 2026; Giri et al., 2025). Based on the literature analysed in this study, both systems converge on the same need for principles to be translated into operational practice (Dinc and Ardic, 2026), thereby showing that the fragmentation is structural rather than jurisdictional.

Accountability becomes progressively distributed and more difficult to trace through AI lifecycle stages. Continuous post-deployment monitoring is often suggested but rarely carried out in practice (Bai et al., 2026; Cary et al., 2025). Organisational and professional responsibilities remain undefined (Abu Assab et al., 2026; Wilkie et al., 2025) and only six of eighty papers discussed legal liability (Hoffman, 2021; Savastano et al., 2025). No explicit framework for allocating responsibility or compensating patients was identified within the included literature, creating a gap between organisational governance before deployment and mechanisms for addressing post-deployment algorithmic harm.

4.4 Strengths and Limitations of the Review

The search covered only Scopus and English-language publications, so relevant work indexed elsewhere or in other languages may have been missed, particularly from low and middle income countries. Screening relied on structured evidence mapping and weighted relevance ranking rather than duplicate independent full-text review. As screening was conducted by one reviewer, some selection bias cannot be excluded. Papers using different terminology or addressing equity implicitly may have been missed. Publication type contributed a source-type score to the relevance ranking rather than constituting a formal appraisal of methodological quality.

These limitations are balanced by clear strengths. The review followed PRISMA-ScR framework with a documented and transparent screening and ranking process, and synthesises 80 studies from 2016 to 2026 across diverse designs and regions.

4.5 Implications and Future Research

For developers, equity cannot be restricted to dataset diversification and fairness metrics. It must be built across the lifecycle and treated as a foundational design principle rather than an evaluation afterthought.

For healthcare organisations, implementation is itself an equity intervention: workflow design, clinician training and local monitoring determine whether a fair model delivers fair care. For policymakers, the priority is connecting regulatory instruments to organisational practice. Across all three, underserved communities must be repositioned from recipients of AI to participants in its design and governance.

The evidence gaps summarised in Table 4 define a future research agenda. Disability, digital exclusion and legal liability are underrepresented in the final corpus despite their importance for fairness and there is no long-term evidence on whether these systems remain fair after implementation. Future research should prioritise accessible and inclusive CDMS design for disabled patients, equity evaluation in digitally excluded and low-resource settings, intersectional equity assessment, and empirical study of how AI-mediated systems perform across populations after routine adoption.

Table 5. Research gaps and future research priorities.

Current state	*Identified gap*	*Future research priority*
Algorithmic fairness	***Focus on technical metrics***	***Socio-technical evaluation***
Responsible AI	***Principle-based***	***Operational implementation***
Accessibility	***Interface-focused***	***Whole-lifecycle inclusion***

Governance	***Fragmented***	***Integrated governance***
Health equity	***Population-specific***	***Cross-population synthesis***

5. Conclusion

Artificial intelligence is transforming clinical decision-making. It offers potential to improve diagnostic accuracy, clinical efficiency and healthcare delivery. This review demonstrates that these opportunities are not distributed equitably. Health inequities associated with AI-mediated CDMS occur not only from algorithmic bias but also from the interaction between technical design, healthcare infrastructure, organisational practices and wider social determinants that disproportionately affect underserved communities.

Across the reviewed literature, racialised communities, disabled patients and digitally excluded populations remain at increased risk of being excluded from AI-enabled healthcare. Despite progress in addressing fairness, transparency and explainability, these efforts remain largely fragmented. Combining evidence from design, law and healthcare, this review shows that equitable AI cannot be achieved through siloed technical interventions or regulatory measures alone. Approaches recognising the interdependence of design, governance, organisational practice and patient experience must be considered throughout the AI lifecycle.

While Responsible Design is broadly recognised as an essential perspective to develop equitable AI-mediated CDMS, its application remains fragmented, with limited integration across design, governance and post-deployment practice. Few studies integrate participatory design, inclusive development, algorithmic fairness, and

trustworthy AI into a coherent, holistic governance approach extending beyond development to implementation, monitoring and continuous evaluation (Bokolo et al., 2026; Wilkie et al., 2025).

This review reframes health equity in clinical AI as an ongoing responsibility embedded throughout the socio-technical lifecycle, rather than as a property determined solely by technical performance. It positions Responsible Design as an integrating perspective through which equity, participation, governance and accountability can be considered across the lifecycle. The next critical step for researchers is to move from simply diagnosing bias to designing real-world solutions by testing practical ways to embed equity and public participation into daily clinical practice. The next generation of clinical AI represents a critical turning point: it may either contribute to reducing long-standing gaps in care or reinforce them if equity is not embedded in design, implementation and governance.

Translating this agenda into real-world practice requires deeper collaboration between scholars, regulatory bodies, healthcare providers, researchers, and most importantly patients, community representatives and advocacy organisations representing racialised, disabled and digitally excluded populations.

Acknowledgements

The authors acknowledge that this study was conducted without external funding.

Author Contributions

ID developed the legal underpinnings of the research idea prior to the commencement of his PhD project. He led on research conceptualisation, investigation, literature search, data extraction and charting, analysis and thematic synthesis, as well as original manuscript drafting. FC provided input on the methodological framing and on the alignment of the discussion with the scope of the journal, offered critical intellectual input and supervised the project. All authors approved the manuscript before submission and agreed to be accountable for the work.

REFERENCES

Abu Assab, M., Al Bahar, F., Abu Dayyih, W., Alazazmeh, B. M., Assaf, S. W., Abed, A., Alrasheed, H. A., & Zakaraya, Z. (2026). Healthcare professionals' perceptions of AI-assisted clinical decision-making in Jordan: A qualitative study of trust, accountability, system readiness, and professional practice. Healthcare. https://doi.org/10.3390/healthcare14121724

Adegunle, F., Chhatwal, K., Arab, S., Alabdaljabar, M. S., Raslan, M. A., Sayed, O., & Goldsweig, A. M. (2026). Bias and oversight in clinical AI: A review of decision support tools and equity frameworks. Journal of General Internal Medicine. https://doi.org/10.1007/s11606-026-10229-5

Al-Duais, F. S., Lababneh, A. S. K., Smerat, A., & Taloba, A. I. (2026). Fairness-aware and trustworthy machine learning framework for secure stroke risk prediction in privacy-sensitive clinical data environments. Journal of Internet Services and Information Security. https://doi.org/10.58346/JISIS.2026.I2.034

Ali, M., Ali, S., Abbas, Q., Abbas, Z., & Lee, S. W. (2025). Artificial intelligence for mental health: A narrative review of applications, challenges, and future directions in digital health. Digital Health. https://doi.org/10.1177/20552076251395548

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19–32. https://doi.org/10.1080/1364557032000119616

Arockiasamy, J. M. I. (2025). Responsible AI for global health equity: A decentralized framework. International Symposium on Technology

and Society (ISTAS) Proceedings. https://doi.org/10.1109/ISTAS65609.2025.11269620

Bagali, M. M., SaiGanesh, S., Sharma, S., & Alkhayya, A. (2026). ETHIC-AID: A framework for ethical AI deployment in healthcare with bias mitigation, transparency, and accountability controls. In Medical minds and machine learning: Transforming care with AI. https://doi.org/10.1201/9781003726913-9

Bai, B., Liu, X., & Li, H. (2026). Federated multimodal AI for precision-equitable diabetes care. Frontiers in Digital Health. https://doi.org/10.3389/fdgth.2025.1678047

Baruah, A., & Majumdar, I. G. (2026). Transforming dementia care in India: The promises and challenges of AI and digital health technologies. In Digital technologies and aging in South Asia and Middle East: Transforming elderly care. https://doi.org/10.4324/9781003740575-19

Bashir, S. G., Salad, H. A., Abdullahi, Y. B., Abdi, Y. H., Abdi, M. S., Ahmed, N. I., & Saidu Musa, S. (2026). Artificial intelligence approaches to predicting treatment non-adherence in chronic diseases: A narrative review. Frontiers in Digital Health. https://doi.org/10.3389/fdgth.2026.1769337

Belmejdoub, N., Rektouti, M., Mouniane, Y., Ouardi, B. E., Ouaaziz, N. A., Elkhlifi, F., Njili, S., Hmouni, D., Bour, A., Khal-Layoun, S., Ouhssine, M., Benkayba, W., Belomaria, M., & Bakkali, M. E. (2026). Algorithmic neutrality in healthcare AI: Critical perspectives on bias, inequality, and the (re)production of health disparities. In AI and digital innovation in healthcare reform: From policy to practice. https://doi.org/10.4018/979-8-2600-0939-0.ch008

Blessing, L. T., & Chakrabarti, A. (2009). DRM, a design research methodology. London: Springer London.

Bokolo, A. J., Denecke, K., & Gabarron, E. (2026). A patient-centered approach to inclusive AI in healthcare: Potential benefits, challenges, and recommendations. Health and Technology. https://doi.org/10.1007/s12553-026-01080-z

Braveman, P., & Gruskin, S. (2003). Defining equity in health. Journal of epidemiology and community health, 57(4), 254-258.

Bray, N. N., Wheeler, D., & Zumwalt, J. (2025). AI-enhanced rural medical education: Bridging gaps and building trust. Academic Medicine. https://doi.org/10.1097/ACM.0000000000006105

Budhu, J. A., Anderson, N., Branson, C., Dy-Hollins, M. E., Jimenez-Gomez, A., Nearing, K., Saleh Velez, F. G., Staisch, J., Wabulya, A., & Mendizabal, A. (2025). Health equity considerations in the age of artificial intelligence. Neurology. https://doi.org/10.1212/WNL.0000000000214356

Cappello, J., El Morr, C., & Dolatabadi, E. (2026). Aligning AI for health equity: Successes, gaps, and barriers in the Canadian healthcare system. In AI for a just world: Power, liberation, and the people left behind. https://doi.org/10.1201/9781003737810-5

Cary, M. P., Bessias, S., McCall, J., Pencina, M. J., Grady, S. D., Lytle, K., & Economou-Zavlanos, N. J. (2025). Empowering nurses to champion health equity & BE FAIR: Bias elimination for fair and responsible AI in healthcare. Journal of Nursing Scholarship. https://doi.org/10.1111/jnu.13007

Chandra, J., Kaur, R., & Sahay, R. (2025). Integrated framework for equitable healthcare AI: Bias mitigation, community participation, and regulatory governance. 2025 IEEE 14th International Conference

on Communication Systems and Network Technologies (CSNT). https://doi.org/10.1109/CSNT64827.2025.10968102

Cheong, P. L., & Kubota, K. (2026). Advancing global nursing policy: AI-driven health literacy innovations for migrant domestic workers. International Nursing Review. https://doi.org/10.1111/inr.70149

Chin, M. K., Đoàn, L. N., Russo, R. G., Roberts, T., Persaud, S., Huang, E., Fu, L., Kui, K. Y., Kwon, S. C., & Yi, S. S. (2023). Methods for retrospectively improving race/ethnicity data quality: A scoping review. Epidemiologic Reviews. https://doi.org/10.1093/epirev/mxad002

Clifford, G., Nguyen, T., Shaw, C., Newton, B., Francis, S., Salari, M., Evans, C., Jones, C., Akintobi, T. H., & Taylor, H. (2022). An open-source privacy-preserving large-scale mobile framework for cardiovascular health monitoring and intervention planning with an urban African American population of young adults: User-centered design approach. JMIR Formative Research. https://doi.org/10.2196/25444

Das, D. C., Rony, M. K. K., Dutta, S., Zujbe, M. D. S. A. Z., Bhattacharjee, T., Debnath, N., Maruf, S. A., Mortuza, C. G., Rahman, F., & Das, A. (2026). The role of artificial intelligence in medication management for older adults: A systematic review. Aging Medicine. https://doi.org/10.1002/agm2.70074

Denecke, K., Vagliano, I., Hewitt, L., Altamimi, A.-K., Meystre, S., & Teodoro, D. (2026). Ethical implications of small language models (SLMs) in healthcare applications. Studies in Health Technology and Informatics. https://doi.org/10.3233/SHTI260236

Dinc, R., & Ardic, N. (2026). Deploying AI for global health equity: A readiness matrix and case studies from low- and middle-income

countries. Health Informatics Journal. https://doi.org/10.1177/14604582261455105

Ekuta, V. (2025). Racing against the algorithm: Leveraging inclusive AI as an antiracist tool for brain health. Clinical and Translational Science. https://doi.org/10.1111/cts.70364

Fesabas, A. K. R., & Shaker, V. (2025). Exploring ethical AI in digital mental health therapy: Accountability, transparency, and empowerment in Southeast Asia. Annual Review of CyberTherapy and Telemedicine.

Giri, P., Katole, A., & Jha, A. (2025). The role of artificial intelligence in mental health: A comprehensive review. Indian Journal of Community Health. https://doi.org/10.47203/IJCH.2025.v37i04.003

Gupta, S., & Jaiswal, R. (2025). How can we improve the insurance industry decision-support system: Integrating people's perspective. Information Discovery and Delivery. https://doi.org/10.1108/IDD-09-2024-0142

Hasanzadeh, F., Josephson, C. B., Waters, G., Adedinsewo, D., Azizi, Z., & White, J. A. (2025). Bias recognition and mitigation strategies in artificial intelligence healthcare applications. NPJ Digital Medicine, 8(1), 154.

Hilling, D. E., Ihaddouchen, I., Buijsman, S., Townsend, R., Gommers, D., & van Genderen, M. E. (2025). The imperative of diversity and equity for the adoption of responsible AI in healthcare. Frontiers in Artificial Intelligence. https://doi.org/10.3389/frai.2025.1577529

Hoffman, S. (2021). The emerging hazard of AI-related health care discrimination. Hastings Center Report. https://doi.org/10.1002/hast.1203

Husebo, B. S., Haugland‐Pruitt, J., & Patrascu, M. (2025). Ethical challenges for artificial intelligence in clinical practice and research: A case study on people with dementia. Neuropsychology in the Age of Digital Health: Exploration, Application, and Practical Implementation, 297-320.

Iloanusi, N.-J., & Chun, S. A. (2024). AI impact on health equity for marginalized, racial, and ethnic minorities. ACM International Conference Proceeding Series. https://doi.org/10.1145/3657054.3657152

Ingraham, N. E., Jones, E. K., King, S., Dries, J., Phillips, M., Loftus, T., Evans, H. L., Melton, G. B., & Tignanelli, C. J. (2023). Re-aiming equity evaluation in clinical decision support: A scoping review of equity assessments in surgical decision support systems. Annals of Surgery. https://doi.org/10.1097/SLA.0000000000005661

Kannadasan, K. (2026). Multimodal AI and telemedicine in remote health monitoring. In Multimodal AI in healthcare: Fusion intelligence for advanced health diagnostics. https://doi.org/10.1007/978-981-95-9027-8_12

Kukkar, A., Loibor, J. M., Cifci, M. A., Barman, M. P., Hussain, S., Gaftandzhieva, S., & Doneva, R. (2026). Improved prediction of childhood anemia using hybrid ensemble learning and dual-level explainability. Journal of Public Health Research. https://doi.org/10.1177/22799036261462908

Lacmanovic, S., & Skare, M. (2025). Artificial intelligence bias auditing – current approaches, challenges and lessons from practice. Review of Accounting and Finance. https://doi.org/10.1108/RAF-01-2025-0006

Lan, G., Zhu, Y., Xiao, S., Iqbal, M., & Yang, J. (2025). Mitigating data bias in healthcare AI with self-supervised standardization. IEEE Journal of Biomedical and Health Informatics. https://doi.org/10.1109/JBHI.2025.3588196

Lau, P. L. (2025). Rewriting the narrative of AI bias: A data feminist critique of algorithmic inequalities in healthcare. Law, Technology and Humans. https://doi.org/10.5204/lthj.3845

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. Implementation Science, 5, Article 69. https://doi.org/10.1186/1748-5908-5-69

Liu, D., Lin, Y., Yan, R., Wang, Z., Bold, D., & Hu, X. (2025). Leveraging artificial intelligence for digital symptom management in oncology: The development of CRCWeb. JMIR Cancer. https://doi.org/10.2196/68516

Marvin, G., Jjingo, D., Nakatumba-Nabende, J., & Alam, M. G. R. (2023). Local interpretable model-agnostic explanations for online maternal healthcare. 2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN). https://doi.org/10.1109/ICSTSN57873.2023.10151520

McGrath, C., Chau, C. W. R., & Molina, G. F. (2025). Monitoring oral health remotely: Ethical considerations when using AI among vulnerable populations. Frontiers in Oral Health. https://doi.org/10.3389/froh.2025.1587630

Mendez-Lopez, A., Park, K., & Hensman, R. D. (2026). Changes needed in health information systems to support an ageing population. Bulletin of the World Health Organization. https://doi.org/10.2471/BLT.25.294184

Mohanty, M., & Mohanty, A. (2025). Machine learning-based recommender systems for cancer patients. In Artificial intelligence in oncology: Cancer diagnosis and treatment, medical imaging, and personalized medicine. https://doi.org/10.1007/978-3-031-94302-7_7

Murphy, A., Bowen, K., Naqa, I. M. E., Yoga, B., & Green, B. L. (2025). Bridging health disparities in the data-driven world of artificial intelligence: A narrative review. Journal of Racial and Ethnic Health Disparities. https://doi.org/10.1007/s40615-024-02057-2

Murugesan, G. S., & Viswanathan, S. A. (2024). Enhancing human-machine interaction: A study on deployment of LLM and Gen AI hybrid models in responsible humanoids for human assistance. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). https://doi.org/10.1109/ICCCNT61001.2024.10724819

Mwogosi, A. (2025). Leveraging AI to enhance healthcare delivery in Tanzania: Innovations and ethical imperatives. SAGE Open. https://doi.org/10.1177/21582440251378162

Nair, P., Dai, M., Patel, D. A., Pearlman, K., & Shah, S. D. (2025). Equity by design – Using digital technology to overcome cardiovascular health disparities. Current Cardiology Reports. https://doi.org/10.1007/s11886-025-02319-3

Nouis, S. C. E., Uren, V., & Jariwala, S. (2025). Evaluating accountability, transparency, and bias in AI-assisted healthcare decision-making: A qualitative study of healthcare professionals' perspectives in the UK. BMC Medical Ethics. https://doi.org/10.1186/s12910-025-01243-z

Olawade, D. B., Bolarinwa, O. A., Adebisi, Y. A., & Shongwe, S. (2025). The role of artificial intelligence in enhancing healthcare for people with disabilities. Social Science & Medicine. https://doi.org/10.1016/j.socscimed.2024.117560

Osonuga, A., Osonuga, A. A., Fidelis, S. C., Osonuga, G. C., Juckes, J., & Olawade, D. B. (2025). Bridging the digital divide: Artificial intelligence as a catalyst for health equity in primary care settings. International Journal of Medical Informatics. https://doi.org/10.1016/j.ijmedinf.2025.106051

Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. JBI Evidence Synthesis, 18(10), 2119–2126. https://doi.org/10.11124/JBIES-20-00167

Pillay, T. S., Khan, A. I., & Yenice, S. (2025). Artificial intelligence (AI) in point-of-care testing. Clinica Chimica Acta. https://doi.org/10.1016/j.cca.2025.120341

Putri, V. C. C. (2026). AI-generated visual disinformation and digital equity: An intersectional analysis of algorithmic vulnerabilities among marginalised communities. Frontiers in Human Dynamics. https://doi.org/10.3389/fhumd.2026.1791655

Racine, E. E. (2025). Que(e)rying artificial intelligence use for infectious disease surveillance: The need for a reparative algorithmic praxis. Big Data & Society. https://doi.org/10.1177/20539517241289440

Radanliev, P. (2025). Privacy, ethics, transparency, and accountability in AI systems for wearable devices. Frontiers in Digital Health. https://doi.org/10.3389/fdgth.2025.1431246

Ricchezze, G., Tomassini, L., & Bacci, M. (2025). Ethical reflections on integrating artificial intelligence in care practices. Community Health Equity Research & Policy. https://doi.org/10.1177/2752535X251370928

Roy, M., Singh, M., & Jamwal, M. (2025). Adoption of AI-based diabetes diagnosis: A patient and institutional perspective in Maharashtra and Karnataka. International Journal of Pharmaceutical and Healthcare Marketing. https://doi.org/10.1108/IJPHM-10-2024-0111

Sarkar, S., Das, S., Chanda, A., & Biswas, S. (2025). Classify chronic wounds: The need of explainable AI and responsible AI. In Explainable and responsible artificial intelligence in healthcare. https://doi.org/10.1002/9781394302444.ch12

Savastano, M. C., Rizzo, C., Fossataro, C., Bacherini, D., Giansanti, F., Savastano, A., Arcuri, G., Rizzo, S., & Faraldi, F. (2025). Artificial intelligence in ophthalmology: Progress, challenges, and ethical implications. Progress in Retinal and Eye Research. https://doi.org/10.1016/j.preteyeres.2025.101374

Shukla, V., & Parker, G. G. (2025). On addressing bias and fairness in large language models for responsible innovation in gender contexts. 2025 Gender and Technology Conference (GTC). https://doi.org/10.1109/GTC64325.2025.11478234

Stipelman, C. H., Kukhareva, P. V., Trepman, E., Nguyen, Q.-T., Valdez, L., Kenost, C., Hightower, M., & Kawamoto, K. (2022). Electronic health record-integrated clinical decision support for clinicians serving populations facing health care disparities: Literature review. Yearbook of Medical Informatics. https://doi.org/10.1055/s-0042-1742518

Thorat, P., & Kalani, R. (2026). Artificial intelligence and diversity and inclusion strategies. In AI dynamics in HRM: Revolutionizing practices for the future workplace. https://doi.org/10.1201/9781779641373-5

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Annals of Internal Medicine, 169(7), 467–473. https://doi.org/10.7326/M18-0850

Tsai, T. C., Arik, S., Jacobson, B. H., Yoon, J., Yoder, N., Sava, D., Mitchell, M., Graham, G., & Pfister, T. (2022). Algorithmic fairness in pandemic forecasting: Lessons from COVID-19. npj Digital Medicine. https://doi.org/10.1038/s41746-022-00602-z

Wen, J., Zhang, T., Ye, S., Li, C., Han, R., Huang, R., Shen, B., Chen, A., & Li, Q. (2024). Development of transient ischemic attack risk prediction model suitable for initializing a learning health system unit using electronic medical records. BMC Medical Informatics and Decision Making. https://doi.org/10.1186/s12911-024-02767-x

Whitehead, M. (1992). The concepts and principles of equity and health. International journal of health services, 22(3), 429-445.

Whitehead, M., Dahlgren, G., & World Health Organization. (2006). Levelling up (part 1): a discussion paper on concepts and principles for tackling social inequities in health.

Wilkie, P. Y., Namdar, K., Samsel, K., Pasricha, S. V., Jaiswal, N., Salo, Z., So, M., Panchal, S., Malhotra, A. K., Nashnoush, E., Walker, S. L., Lee, S. M., Mamdani, M., & Celi, L. A. (2025). Bridging accountability

and innovation: Key findings from the Health AI Systems Thinking for Community (HASTC) workshop. Facets. https://doi.org/10.1139/facets-2025-0117

Wong, E., Bermudez-Cañete, A., Campbell, M. J., & Rhew, D. C. (2025). Bridging the digital divide: A practical roadmap for deploying medical artificial intelligence technologies in low-resource settings. Population Health Management. https://doi.org/10.1089/pop.2024.0222

World Health Organization. (2025). World report on social determinants of health equity: executive summary.

Yu, E., Rosenberg, G. M., Udelsman, B. V., Harano, T., Atay, S. M., Kim, A. W., Shakhsheer, B. A., & Wightman, S. C. (2026). Narrative review of the ethics of artificial intelligence: Are we ready for artificial intelligence in surgery? Journal of Thoracic Disease. https://doi.org/10.21037/jtd-2025-1814

Zarnegar, A. (2026). Addressing bias and ensuring fairness in AI systems for healthcare. Proceedings of 2025 18th Conference on Health Informatics Knowledge Management (HIKM). https://doi.org/10.1145/3774816.3774817

APPENDIX

I. NOTE ON METHODOLOGICAL TRANSPARENCY

Manuscript: "Towards Responsible Design for Equitable AI-Mediated Clinical Decision-Making Systems: A Scoping Review of Health Equity and Underserved Populations"

Design for All - Special Issue on Responsible Design: Inclusive, Ethical and Sustainable Futures

The Methods section of this manuscript summarises the search strategy at corpus level, as is conventional for a scoping review of this length. Four features of the strategy warrant explicit statement, and are documented in full in II section 1.

1. The figure of 2,082 records is a composite retrieval total.

Sixteen separate searches were executed in Scopus between early May and June 2026, using an iterative strategy in which each search paired AI and clinical decision support terminology with a specific equity, harm, design or governance concept. The searches were designed to overlap. The figure of 2,082 is therefore the sum of the sixteen individual query yields rather than a count of distinct publications; 385 duplicates were subsequently removed, leaving the 1,697 unique records that entered screening. Per-search yields are reported in Section 2 and the full reconciliation in Section 4.

2. Search parameters were not uniform across the sixteen searches.

Search parameters were not uniform across the sixteen searches. Six searched the full 2016-2026 period, and the remainder used later starting years or an upper limit of 2025. Four searched with an explicit English language limit, and seven restricted document type for

articles and reviews. These variations were a result of the iterative search strategy adopted, while the overall eligibility criteria were applied during screening. Coverage of the governance, liability and automation-bias themes nonetheless extends over a shorter period than coverage of health equity and algorithmic bias.

3. Retrieval for the United Kingdom and India was deliberately targeted.

Three searches restricted retrieval to the United Kingdom or India, contributing 625 records in total, and a fourth targeted low-resource and Global South settings. The prominence of these two systems in the corpus, and the UK–India comparison developed in the Discussion, are therefore a designed contrast between a high-regulatory-maturity system and a rapidly scaling one, rather than an unprompted pattern in the wider literature. The comparison is offered on that basis and the authors do not claim it is representative of the global distribution of research on this topic.

4. Sparse representation of some themes reflects the prioritisation strategy as well as the literature.

Of the 1,553 records assessed against the full inclusion criteria, 857 passed the 15.0-point weighted relevance threshold but were excluded solely for falling below the top-80 relevance cut-off. Four searches targeting liability and accountability returned 677 records in combination, while six of the eighty included studies were coded to liability; disability terms appeared in eight searches and eight included studies were coded to disability. Retrieval volume is not equivalent to topical coverage, and several of these searches overlapped substantially. Nonetheless, where the manuscript describes disability, digital exclusion and legal liability as under

examined, that observation is properly understood as describing their representation within this corpus and the review's prioritisation, rather than as a measured claim about the volume of available literature.

II. Scopus Search Strategy and Record results

Supplementary to: "Towards Responsible Design for Equitable AI-Mediated Clinical Decision-Making Systems: A Scoping Review of Health Equity and Underserved Populations"

Design for All — Special Issue on Responsible Design: Inclusive, Ethical and Sustainable Futures

Section 1 - Search process:

Sixteen searches were executed in Scopus using an iterative strategy. Each search combined terms for Artificial Intelligence (AI)-mediated Clinical Decision-Making Systems (CDMS) with a specific equity, harm, design or governance concept, and successive searches were refined in response to the records retrieved. Search parameters therefore varied across the individual searches.

The sixteen searches returned 2,082 records in total. Because the searches were designed to overlap, this figure is the sum of individual query results rather than a count of distinct publications. Following export into the review database, 385 duplicates were removed (365 matched by Digital Object Identifier (DOI) and 20 identified through title and publication-year matching) leaving 1,697 unique publications. This dataset formed the basis for title and abstract screening, assessment against the full inclusion criteria, and structured relevance scoring across seven conceptual domains.

Boolean operators, phrase searching, single-character wildcards (?) and truncation (*) were used throughout. The wildcard form

raciali?ed retrieves both British and American spellings. Search parameters were not uniform across the sixteen searches; the variation is documented in Section 2 and its implications are set out in Section 4.

Section 2 - Search characteristics and yields

	Focus	*Years*	*English*	*Geographic limit*	*n*
1	*Exploratory scoping search — harm and racialised communities*	*2016–2026*	*No*	*—*	*7*
2	*AI, harm and underserved populations — UK and India*	*2016–2026*	*Yes*	*UK / India*	*255*
3	*AI and equity — India-focused*	*2016–2026*	*Yes*	*India*	*362*
4	*Responsible Design and equity — UK and India*	*2021–2026*	*Yes*	*UK / India*	*8*
5	*Harm, liability and accountability with equity*	*2021–2026*	*No*	*—*	*292*
6	*Harm, liability and accountability with equity — widened*	*2016–2026*	*No*	*—*	*329*
7	*Responsible Design and equity — no geographic limit*	*2016–2026*	*Yes*	*—*	*178*
8	*Algorithmic harm and safety*	*2017–2025*	*No*	*—*	*131*
9	*AI-mediated healthcare and patient-centred harm*	*2016–2026*	*No*	*—*	*76*
10	*Automation bias and human oversight*	*2017 onwards*	*No*	*—*	*80*
11	*Participatory design and AI governance*	*2019 onwards*	*No*	*—*	*188*
12	*Cross-jurisdictional regulation and low-resource settings*	*2020 onwards*	*No*	*LMIC / Global South*	*17*

13	*Responsible Design and governance mechanisms*	*2019 onwards*	*No*	*—*	*62*
14	*Algorithmic harm, legal frameworks and compensation*	*2018 onwards*	*No*	*—*	*25*
15	*Training-data bias in diagnosis and triage*	*2019 onwards*	*No*	*—*	*41*
16	*CDSS harm, liability and oversight*	*2019 onwards*	*No*	*—*	*31*
				Total:	*2,082*

Note: The searches used varied retrieval parameters as part of the iterative strategy. Six of the sixteen searches used the full 2016–2026 window; four applied an explicit English-language limit; and seven restricted document type to articles and reviews. The overall eligibility criteria were applied during screening.

Section 3 - Search queries as executed

Reproduced verbatim. Record counts are those returned at the time of searching; Scopus coverage changes over time and re-running these queries may return different totals.

Search 1. Exploratory scoping search — harm and racialised communities n = 7

TITLE-ABS-KEY (healthcare AND harm AND ("racialized communities" OR "affected populations"))

AND PUBYEAR > 2015 AND PUBYEAR < 2027

Search 2. AI, harm and underserved populations — UK and India n = 255

TITLE-ABS-KEY (("artificial intelligence" OR AI OR "machine learning" OR "clinical decision

support system*" OR "clinical decision-making" OR "AI-mediated" OR algorithm*) AND (healthcare

OR health OR medicine OR clinical OR patient*) AND (harm OR injur* OR "adverse event*" OR

"patient safety" OR inequity OR inequit* OR unfair* OR fairness OR bias OR discrimination OR

"health disparit*" OR "health inequit*") AND ("raciali?ed communit*" OR "raciali?ed

population*" OR minorit* OR underserved OR vulnerab* OR marginali?ed OR disability OR disabled

OR "digital exclusion" OR "digital divide" OR "digital literacy" OR "social exclusion") AND (

"United Kingdom" OR UK OR England OR Scotland OR Wales OR "Northern Ireland" OR India)) AND

PUBYEAR > 2015 AND PUBYEAR < 2027 AND (LIMIT-TO (LANGUAGE , "English"))

Search 3. AI and equity — India-focused n = 362

TITLE-ABS-KEY (("artificial intelligence" OR AI OR "machine learning" OR "clinical decision

support" OR algorithm*) AND (healthcare OR health OR medicine) AND (equity OR fairness OR

disability OR bias OR "digital exclusion" OR underserved) AND (India OR "Indian healthcare" OR

"National Digital Health Mission" OR Ayushman)) AND PUBYEAR > 2015 AND PUBYEAR < 2027 AND (

LIMIT-TO (LANGUAGE , "English"))

Search 4. Responsible Design and equity — UK and India n = 8

TITLE-ABS-KEY (("artificial intelligence" OR AI OR "machine learning" OR "clinical decision

support system*" OR "clinical decision support" OR "clinical decision-making" OR "AI-enabled

healthcare" OR "AI-mediated healthcare" OR algorithm*) AND (healthcare OR health OR medicine

OR clinical OR patient*) AND (equity OR fairness OR inclusion OR inequit* OR disparit* OR

discrimination OR bias OR accessibility OR "health equity" OR "healthcare disparities") AND (

"responsible design" OR "inclusive design" OR "human-centred design" OR "human centered design"

OR "universal design" OR "participatory design" OR co-design OR "responsible innovation" OR

"responsible AI" OR "trustworthy AI") AND (disabled OR disability OR "raciali?ed communit*" OR

"raciali?ed population*" OR minority OR vulnerable OR underserved OR marginali?ed OR "digital

exclusion" OR "digital literacy" OR "social inequality" OR "health inequality") AND ("United

Kingdom" OR UK OR Britain OR England OR Scotland OR Wales OR "Northern Ireland" OR India)) AND

PUBYEAR > 2020 AND PUBYEAR < 2027 AND (LIMIT-TO (LANGUAGE , "English"))

Search 5. Harm, liability and accountability with equity n = 292

TITLE-ABS-KEY (("artificial intelligence" OR AI OR "machine learning" OR "clinical decision

support system*" OR "clinical decision-making") AND (harm OR injury OR "adverse event*" OR

liability OR accountability) AND (equity OR fairness OR inequit* OR disparit* OR

discrimination OR bias) AND ("racialised communit*" OR disability OR disabled OR "digital

exclusion" OR "digital literacy" OR "underserved population*" OR "marginalised population*" OR

"disadvantaged population*")) AND PUBYEAR > 2020 AND PUBYEAR < 2027

Search 6. Harm, liability and accountability with equity — widened n = 329

TITLE-ABS-KEY (("artificial intelligence" OR AI OR "machine learning" OR "clinical decision

support system*" OR "clinical decision-making") AND (harm OR injury OR "patient safety" OR

"adverse event*" OR liability OR accountability) AND (equity OR fairness OR inequit* OR

disparit* OR discrimination OR bias) AND ("raciali?ed communit*" OR disability OR disabled OR

"digital exclusion" OR "digital literacy" OR "underserved population*" OR "marginali?ed

population*" OR "disadvantaged population*")) AND PUBYEAR > 2015 AND PUBYEAR < 2027

Search 7. Responsible Design and equity — no geographic limit n = 178

TITLE-ABS-KEY (("artificial intelligence" OR AI OR "machine learning" OR "clinical decision

support system*" OR "clinical decision-making" OR "AI-enabled healthcare" OR "AI-mediated

healthcare" OR algorithm*) AND (healthcare OR health OR medicine OR clinical OR patient*) AND

(equity OR fairness OR inclusion OR inequit* OR disparit* OR discrimination OR bias OR

accessibility OR "health equity" OR "healthcare disparities") AND ("responsible design" OR

"inclusive design" OR "universal design" OR "human-centred design" OR "human centered design" OR

"participatory design" OR co-design OR "design for all" OR "responsible innovation" OR

"responsible AI" OR "trustworthy AI") AND (disabled OR disability OR "raciali?ed communit*" OR

"raciali?ed population*" OR minority OR vulnerable OR underserved OR marginali?ed OR "digital

exclusion" OR "digital literacy" OR "social inequality" OR "health inequality")) AND PUBYEAR >

2015 AND PUBYEAR < 2027 AND (LIMIT-TO (LANGUAGE , "English"))

Search 8. Algorithmic harm and safety n = 131

TITLE-ABS-KEY ((algorithmic harm* OR "clinical decision support") AND (harm OR injury OR

adverse OR safety) AND (disab* OR "racial* communit*" OR minority OR vulnerable OR underserved

OR marginalized OR "digital literacy" OR "digital exclusion")) AND PUBYEAR > 2016 AND PUBYEAR

< 2026

Search 9. AI-mediated healthcare and patient-centred harm n = 76

TITLE-ABS-KEY (("AI-mediated healthcare" OR "clinical decision-support systems" OR "patient-

centred harm") AND (equity OR fairness OR inequality OR disparities OR inclusion OR

discrimination OR bias)) AND ("healthcare disparities" OR "racialised populations" OR

"disadvantaged populations" OR "digitally and socially excluded patients" OR "social inequality"

) AND PUBYEAR > 2015 AND PUBYEAR < 2027

Search 10. Automation bias and human oversight n = 80

TITLE-ABS-KEY ("automation bias" OR "over-reliance" OR "algorithm aversion" OR "human-in-the-

loop" OR "meaningful human oversight") AND TITLE-ABS-KEY ("clinical decision" OR "diagnostic

AI" OR "AI-assisted" OR "AI recommendation") AND TITLE-ABS-KEY ("patient safety" OR "harm" OR

"accountability" OR "governance") AND DOCTYPE (ar OR re) AND PUBYEAR > 2016

Search 11. Participatory design and AI governance n = 188

TITLE-ABS-KEY ("participatory design" OR "co-design" OR "patient involvement" OR "patient-

reported" OR "patient-centred participation") AND TITLE-ABS-KEY ("artificial intelligence" OR

"digital health" OR "AI governance" OR "healthcare algorithm") AND TITLE-ABS-KEY ("governance"

OR "accountability" OR "harm reporting" OR "evidence generation" OR "design principles") AND

NOT TITLE-ABS-KEY ("app development" OR "user interface testing" OR "usability testing") AND

DOCTYPE (ar OR re) AND PUBYEAR > 2018

Search 12. Cross-jurisdictional regulation and low-resource settings n = 17

TITLE-ABS-KEY ("regulatory asymmetry" OR "cross-jurisdictional" OR "comparative AI regulation"

OR "global AI governance" OR "regulatory divergence" OR "regulatory capacity") AND TITLE-ABS-

KEY ("artificial intelligence" OR "digital health" OR "healthcare AI") AND TITLE-ABS-KEY (

"low-income" OR "low-resource" OR "Global South" OR "developing countr*" OR "North Africa" OR

"Sub-Saharan" OR "LMIC") AND DOCTYPE (ar OR re) AND PUBYEAR > 2019

Search 13. Responsible Design and governance mechanisms n = 62

TITLE-ABS-KEY ("responsible design" OR "responsible AI" OR "value-sensitive design" OR "human-

centred design" OR "ethical AI design") AND TITLE-ABS-KEY ("governance mechanism" OR

"governance infrastructure" OR "accountability framework" OR "regulatory compliance") AND

TITLE-ABS-KEY ("healthcare" OR "clinical AI" OR "artificial intelligence" OR "medical

technology") AND NOT TITLE-ABS-KEY ("autonomous vehicle" OR "robotics" OR "smart city" OR

"surveillance capitalism") AND DOCTYPE (ar OR re) AND PUBYEAR > 2018

Search 14. Algorithmic harm, legal frameworks and compensation n = 25

TITLE-ABS-KEY ("digital injury" OR "AI-mediated harm" OR "algorithmic harm" OR "data-driven

harm" OR "AI harm" OR "digital harm") AND TITLE-ABS-KEY ("governance" OR "legal framework" OR

"accountability" OR "compensation" OR "compensable" OR "regulatory gap") AND TITLE-ABS-KEY (

"healthcare" OR "clinical" OR "patient" OR "medical AI") AND DOCTYPE (ar OR re) AND PUBYEAR >

2017

Search 15. Training-data bias in diagnosis and triage n = 41

TITLE-ABS-KEY ("algorithmic bias" OR "training data bias" OR "representational bias" OR

"demographic underrepresentation") AND TITLE-ABS-KEY ("diagnosis" OR "triage" OR "treatment

recommendation" OR "clinical pathway") AND TITLE-ABS-KEY ("harm" OR "health disparity" OR

"inequity" OR "misclassification" OR "false negative") AND NOT TITLE-ABS-KEY ("natural

language processing" OR "text classification" OR "sentiment analysis") AND DOCTYPE (ar OR re)

AND PUBYEAR > 2018

Search 16. CDSS harm, liability and oversight n = 31

TITLE-ABS-KEY ("clinical decision support" OR "AI-CDSS" OR "AI-assisted diagnosis" OR "AI-

assisted clinical decision") AND TITLE-ABS-KEY ("harm" OR "misdiagnosis" OR "adverse outcome"

OR "patient safety incident") AND TITLE-ABS-KEY ("governance" OR "accountability" OR

"liability" OR "oversight" OR "regulatory") AND NOT TITLE-ABS-KEY ("image segmentation" OR

"natural language processing" OR "drug discovery" OR "wearable") AND DOCTYPE (ar OR re) AND

PUBYEAR > 2018

Section 4 - Record counts

Stage	*nb*
Sum of results across the sixteen searches	*2,082*
Duplicates removed — matched by DOI	***365***
Duplicates removed — matched by title and publication year	***20***
Total duplicates removed	***385***
Unique records screened by title and abstract	*1,697*
Excluded at relevance screening (no AI/CDMS or equity relevance in title or abstract)	***144***
Records assessed against full inclusion criteria	***1,553***
Excluded: no equity, underserved-population or responsible-design dimension	***296***
Excluded: below 15.0-point weighted relevance threshold	***320***

Excluded: passed threshold but below top-80 relevance cut-off	***857***
Studies included in the review	*80*
Tier 1 (Core) / Tier 2 (Important) / Tier 3 (Relevant)	*27 / 27 / 26*

Section 5 - Limitations arising from the search design

- **Uneven temporal coverage by theme**

Only six of the sixteen searches applied the full 2016–2026 window. Searches 10 to 16, which targeted automation bias, participatory governance, regulatory divergence, algorithmic harm, training-data bias and CDSS liability, began at 2017 to 2020 and carried no upper bound. Search 8 was capped at 2025. The union of all windows spans 2016 to 2026, so the corpus as a whole covers the stated period, but pre-2019 literature on the governance, liability and automation-bias themes was searched less comprehensively than literature on health equity and algorithmic bias.

- **Deliberate geographic targeting of the United Kingdom and India**

Searches 2, 3 and 4 restricted retrieval to the United Kingdom or India, together contributing 625 records, and Search 12 targeted low-resource and Global South settings. The prominence of the United Kingdom and India in the geographic distribution of included studies, and the UK–India comparison discussed in the Discussion, therefore reflect a deliberate feature of the search strategy rather than an unprompted pattern in the wider literature. This does not invalidate the comparison, and should be read as a designed contrast between a high-regulatory-maturity and a rapidly scaling health system rather than as evidence about the global distribution of research.

- **Relevance cut-off and apparent thematic gaps**

Of the 1,553 records assessed against the full inclusion criteria, 857 passed the 15.0-point weighted relevance threshold but were excluded solely for falling below the top-80 cut-off. Retrieval for liability and accountability was substantial — Searches 5, 6, 14 and 16 returned 677 records in total — yet only six of the eighty included studies were coded to liability. Similarly, disability terms appeared in eight of the sixteen searches, and eight included studies were coded to disability. The sparse representation of these themes in the final corpus therefore reflects the

review's prioritisation strategy as opposed to the volume of available literature, and claims about under examination should be read with that in mind.

- **Single database and lack of independent validation**

Retrieval was limited to Scopus, so relevant work indexed only in PubMed, Web of Science, IEEE Xplore, ACM Digital Library or regional databases was not captured. The search strategy was not validated by an independent information specialist. Screening was carried out by a single reviewer in consultation with another member of the study team. Grey literature, preprints and non-English publications were not searched.