



Structural Inequity in Digital Health: Compounding Model of Algorithmic Bias Across the Data-Model-Deployment-Governance Pipeline

Kinjal Nayak  

College of Public Health, Kent State University, United States

Correspondence: Kinjal Nayak, Email: drkinjalnayak@gmail.com

Abstract

Introduction: Digital health technologies including clinical decision-support systems, digital phenotyping, and mental health applications offer population health benefits but risk perpetuating inequalities through algorithmic bias, structural inequities in health system data, and differential digital access. **Methods:** SANRA-guided narrative-review with thematic synthesis was conducted using structured searches across PubMed, Web of Science, and CINAHL for English-language literature published from January 1, 2018, through May 17, 2026. Of 2,304 database records identified, 300 records were retained for title and abstract screening after deduplication and preliminary relevance assessment. Forty sources underwent full-text review, and 22 sources were included in the final narrative synthesis. **Results:** Six interrelated domains were identified: (1) digital pathways, social determinants, and generalizability; (2) algorithmic bias and clinical decision-making; (3) digital phenotyping and mental health inequity; (4) digital access, use, and infrastructure; (5) data foundations and structural bias; and (6) lifecycle governance and accountability structures. Thematic synthesis identified ways in which these inequities may interact across multiple pipeline levels in a manner consistent with a compounding structural process. Based on the synthesis, we propose the Structural Compounding Model of Digital Health Inequity, a four-level conceptual model covering Data, Model, Deployment, and Governance pipeline stages. **Conclusion:** Algorithmic bias and digital health inequity often reflect structural forces including historical underinvestment, biased data collection, proxy outcome selection, and unequal access to digital technologies. Equitable outcomes are likely to require coordinated attention across all four tiers of the Structural Compounding Model of Digital Health Inequity.

Key Words: Algorithmic bias, Digital health equity, Health inequities, Narrative review, Structural compounding model

Introduction

Algorithmic bias refers to systematic errors in computational system outputs that produce inequitable outcomes across demographic groups, arising from biased training data, model design choices, or deployment context.^[1,2] Structural inequity describes the deeply embedded, institutional, and historical forces that sustain unequal distributions of power, resources, and opportunity, including within health care.^[3,4] The term compounding, as used in this review, refers to a proposed process in which bias introduced at one pipeline stage may be reinforced by subsequent stages, producing cumulative inequity across the digital health pipeline. This conceptualization builds on evidence that bias can propagate across stages of the machine-learning pipeline and that interventions targeting only one stage may be insufficient when upstream and downstream conditions remain unaddressed.^[5] Evidence from a commercially deployed clinical risk score demonstrates that algorithmic tools can exhibit racial bias^[1]; algorithm design choices may contribute to health inequities^[2,6];

proxy outcomes used for training can encode structural disadvantage^[7]; bias may accumulate across stages of the machine learning pipeline^[5]; and post-deployment monitoring remains necessary to detect and correct these harms.^[8] Without deliberate attention to equity, algorithmic tools may entrench rather than close existing care gaps.^[1,7,9]

Social and structural conditions are primary determinants of health outcomes.^[3,4] Analytical frameworks that disregard these structural conditions risk misattributing socially produced inequities as individual failings, an error with direct consequences for algorithm design and deployment.^[7,10] Inclusion in digital health has itself been reframed as a social determinant of health, as access, literacy, device ownership, and connectivity can compound socioeconomic disadvantage.^[3,11] Health-related digital inequities may persist even after accounting for some socioeconomic and health-need factors.^[12] Several adjacent frameworks already address digital health equity, AI lifecycle assessment, implementation equity, ethical appraisal, and sources of algorithmic harm.

Suresh and Guttag^[13] catalogue sources of machine-learning harm across the lifecycle; Char DS et al^[14] provide an ethics appraisal framework for machine-learning healthcare applications; HEAAL offers an applied lifecycle equity assessment framework for healthcare delivery organizations adopting AI; and the Digital Health Equity-Focused Implementation Research model links equity-focused digital health and implementation research.^[15,16] Additional lifecycle frameworks, including AI/ML total product lifecycle equity considerations, the evidence- and consensus-based Digital Healthcare Equity Framework, and recent synthesis of telehealth and digital health equity frameworks, further show that lifecycle and implementation approaches to digital health equity are active areas of scholarship.^[17-19] However, these frameworks primarily emphasize assessment, ethical appraisal, implementation guidance, or bias mitigation rather than explicitly theorizing structural inequity as a sequential compounding process across data, model, deployment, and governance levels. The Structural Compounding Model is intended to address this narrower gap by treating structural social disadvantage and digital access inequity as pipeline inputs and by formalizing the proposed compounding mechanism across data, model, deployment, and governance stages.

Two observations highlight the significance of this gap. First, a recent scoping review found that 74.7% of 91 clinical machine learning algorithms exhibited sociodemographic bias; of these, 87% discriminated against socioeconomically disadvantaged groups, and only 1.1% incorporated intersectional analysis.^[20] Second, the European Union Artificial Intelligence Act, which came into effect in August 2024, is highly relevant to healthcare AI and requires conformity assessment for many high-risk AI systems; however, in the authors' interpretation, it does not make demographic subgroup performance reporting or health-equity impact assessment explicit stand-alone conditions for market authorization.^[21] To address this gap, the present review introduces the Structural Compounding Model of Digital Health Inequity.

Methods

Design:

A narrative review design with thematic synthesis was employed. The review was guided by the Scale for the Assessment of Narrative Review Articles (SANRA) to enhance methodological rigor and transparency.^[22] SANRA is a quality appraisal instrument for narrative reviews; it was applied here to assess and strengthen the methodological approach of the review, not as a

reporting guideline in the manner of PRISMA.

Search Strategy:

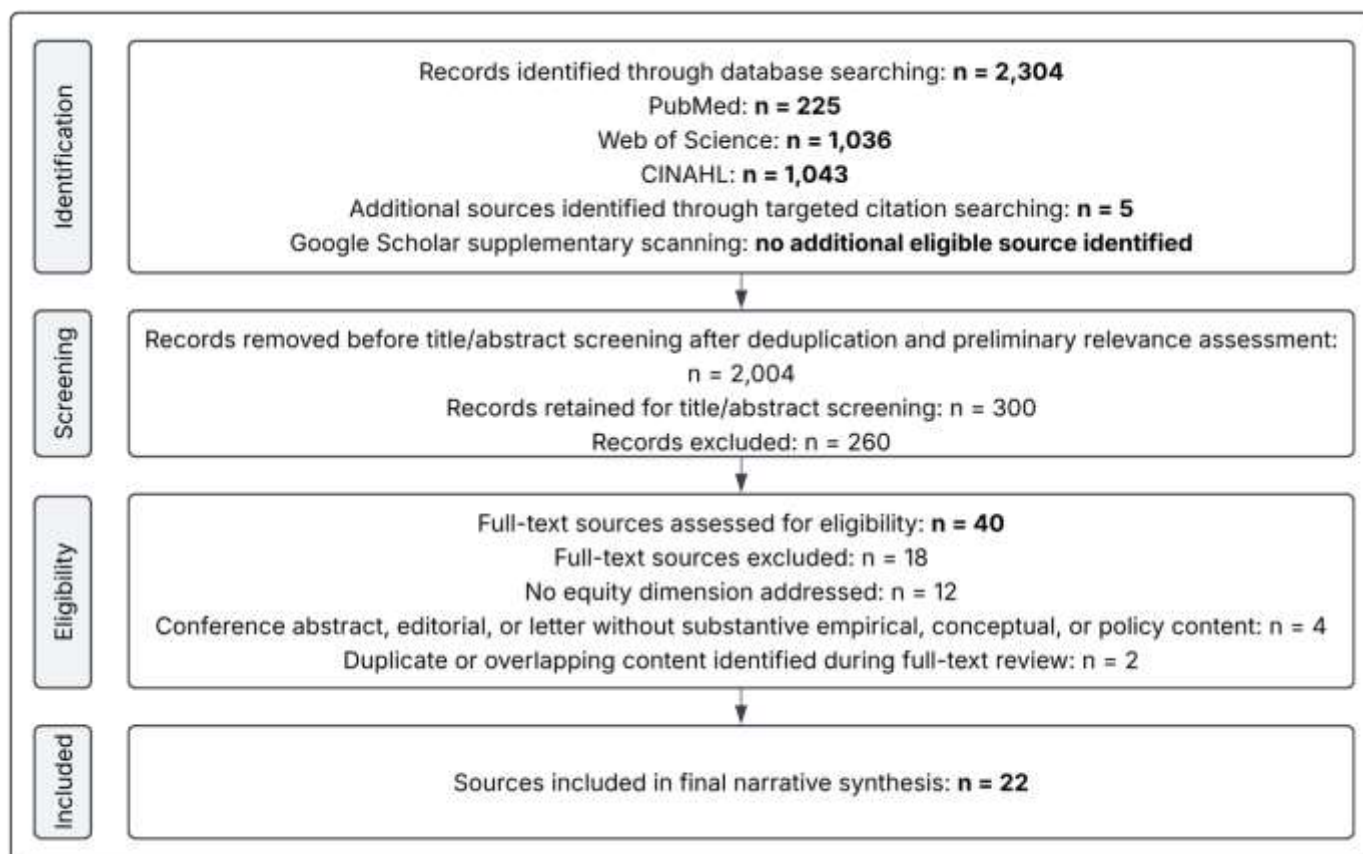
A structured Boolean search strategy was implemented in PubMed, Web of Science, and CINAHL for English-language literature published from January 1, 2018, through May 17, 2026. The 2018 lower bound was selected to capture literature following the emergence of large-scale commercial clinical algorithm deployment and growing policy attention to digital health equity. Search concepts were organized into three thematic blocks: (1) algorithm-related terms (algorithmic bias, algorithmic fairness, machine learning bias, AI bias, clinical algorithm variants, risk prediction models, decision support algorithms); (2) digital health terms (digital health, digital phenotyping, mHealth, mobile health, telehealth, digital biomarkers, passive sensing, wearable devices, patient portals, health apps); and (3) equity and determinants terms (health equity, health disparities, social determinants of health, structural inequity, structural racism, marginalized populations, digital divide, digital literacy). PubMed was searched using MeSH headings (Artificial Intelligence, Machine Learning, Decision Support Systems Clinical, Health Equity, Healthcare Disparities, Social Determinants of Health, Socioeconomic Factors, Racism, Telemedicine, Mobile Applications) combined with Title/Abstract field tags. Equivalent Boolean topic-search syntax was applied to Web of Science and CINAHL using database-appropriate field tags. Google Scholar was used as a supplementary citation-checking and relevance-scanning source; after duplicate removal and relevance assessment, it did not yield any additional eligible sources beyond those identified through the three primary databases.

The full PubMed search string (Supplementary File 1) applied and equivalent Boolean topic-search syntax was applied to Web of Science and CINAHL using the same three thematic blocks with database-appropriate field tags.

Study Selection:

A total of 2,304 records were identified through primary database searching: PubMed (n = 225), Web of Science (n = 1,036), and CINAHL (n = 1,043). Because the search strategy was intentionally broad, 2,004 records were removed before title and abstract screening after deduplication and preliminary relevance assessment. The remaining 300 records were screened by title and abstract, of which 260 were excluded. Forty sources underwent full-text review; these consisted of 35 sources identified through database screening and five sources identified through targeted citation searching because of their direct relevance to framework development. Eighteen full-text sources were excluded

Figure 1:
Source Identification and Selection Summary for the Narrative Review



because they did not address an equity dimension ($n = 12$), were conference abstracts, editorials, or letters without substantive empirical, conceptual, or policy content ($n = 4$), or were duplicate or overlapping records identified during full-text review ($n = 2$). Twenty-two primary sources were included in the final narrative synthesis and are summarized in Table 1. Additional peer-reviewed or authoritative sources were cited to support background, methods, regulatory context, framework comparison, and implications, but these were not counted among the 22 primary synthesis sources. [Figure 1] This review was not prospectively registered. Eligible sources: (a) addressed equity dimensions in digital health data, design, development, analysis, deployment, or governance; (b) examined clinical algorithm fairness, digital phenotyping, digital access, AI/ML bias, or digital health inequity; and (c) were peer-reviewed empirical studies, reviews, conceptual papers, consensus statements, or authoritative regulatory/policy analyses relevant to digital health equity. Authoritative consensus or policy documents were defined as peer-reviewed consensus statements, regulatory analyses, or framework papers produced through transparent expert, stakeholder, policy, or multidisciplinary processes. Sources were excluded if they addressed AI or digital health without an equity dimension, were

conference abstracts, editorials, or letters without substantive empirical, conceptual, or policy content, were non-English publications, were duplicate records, or were outside the review date range.

Quality Assessment:

Because the review included heterogeneous empirical, conceptual, consensus, and regulatory sources, no single formal risk-of-bias tool was applicable across all source types. Instead, included sources were appraised narratively for relevance to digital health equity, clarity of methods or argument, credibility of data source or consensus process, and contribution to framework development. The absence of formal risk-of-bias scoring is acknowledged as a limitation.

Thematic Synthesis:

An inductive coding approach was used: codes were derived from the data rather than imposed from a predetermined framework, allowing themes to emerge from the literature, consistent with established methods for thematic synthesis in narrative reviews.^[23] The primary author extracted each source's study design, population or data source, digital health domain, equity dimension, and relevance to the proposed model. Codes were grouped into candidate themes, compared across source types, and refined into six final domains. Inter-rater agreement could not be assessed because

screening and coding were conducted by a single reviewer.

Six major themes were identified: (1) digital pathways, social determinants, and generalizability; (2) bias in algorithms and clinical decision-making; (3) digital phenotyping and mental health equity; (4) digital access, usage, and infrastructure; (5) data foundations and structural bias; and (6) governance through the lifecycle. Across these six themes, three conceptual layers are evident: a structural environment layer (themes 4 and 5, encompassing access barriers and data foundations), an algorithmic processes layer (themes 1, 2, and 3, covering model bias, phenotyping, and digital pathways), and a governance and mitigation layer (theme 6). This three-layer hierarchy is formalized in the proposed Structural Compounding Model (Table 2).

Results

This section presents synthesized findings from the 22 primary sources included in the narrative synthesis, supported where necessary by supplementary contextual sources used for framework comparison, governance interpretation, and implementation implications. Table 1 provides concise characteristics of the 22 primary sources, while the text below provides thematic interpretation rather than repeating each table entry.

Digital Pathways, Social Determinants, and Generalizability:

Cui^[24] described digital pathways through which data representation bias and distributional shift can reduce AI model generalizability and affect health equity across populations. Mhasawade et al.^[25] argued that incorporating social determinants into machine-learning approaches can support fairness-aware public and population health modelling. Charpignon et al.^[26] further identified physical, biological, human factor, and interpretation biases as mechanisms through which structural conditions may distort clinical measurements before model development begins.

Algorithmic Bias and Clinical Decision-Making:

Obermeyer et al.^[1] reported evidence that a commercially deployed risk stratification algorithm systematically underestimated the care needs of Black patients because it used healthcare costs as a proxy for health need, encoding historical access disparities into its predictions. Benjamin^[10] offered a broader sociological explanation, arguing that quantified metrics can legitimize structural inequalities by lending them scientific authority. Otokiti et al.^[27] highlighted demographic reporting gaps and equity concerns in

clinical AI and machine-learning studies. Hussain et al.^[28] synthesized evidence that AI in healthcare may exacerbate ethnic and racial disparities and emphasized the need for stronger governance and mitigation approaches.

Digital Phenotyping and Mental Health Equity:

Passive sensing through smartphones, including movement, location, activity, communication, and sleep-related signals, is increasingly used to estimate or predict health-related outcomes with minimal patient burden.^[29] Within the sources reviewed, equity-specific evidence in digital phenotyping and digital mental health appeared less empirically developed than evidence on clinical algorithmic decision-making.^[20,29,31] Lee et al.^[29] conducted a scoping review of 40 studies showing that passive smartphone data, including GPS, mobility, activity, sleep, social activity, and in-phone activity, are used to correlate with or predict health-related outcomes. Birk and Samuel^[30] cautioned that digital phenotyping may rely on assumptions of normality, treat digital data as a proxy for social life, and risk reifying or medicalizing mental health problems without sufficient attention to social context. Buda et al.^[31] applied a fairness framework to a commercial digital well-being service, supporting the relevance of fairness assessment in this domain. Rowe et al.^[32] and Robinson et al.^[33] further emphasize the need to embed equity, diversity, and ongoing evaluation into digital mental health development and implementation. The reviewed literature suggests at least three conditions that may be important for digital phenotyping to benefit marginalized populations: sufficient smartphone and data connectivity among vulnerable groups; demographic validation of behaviour-based algorithms; and clinician capacity to distinguish structural from genuinely clinical signals.^[29,30,34] Together, these sources indicate that digital phenotyping may create equity risks when behavioural data are interpreted without social, environmental, and cultural context.

Digital Accessibility, Usage, and Infrastructural Barriers:

Structural barriers to digital health access are consistently documented across population groups. Badr et al.^[11] conducted a scoping review of 41 studies and found that digital inequalities follow consistent sociodemographic patterns, with policy recommendations emphasizing equity-oriented digital strategies. Zhang et al.^[12] used NHS administrative data in a population-level analysis and found that area-level deprivation independently predicted lower use of digital health services after controlling for health need, age, and other sociodemographic factors, demonstrating a digital access gradient that is not fully explained by the

Table 1:
Primary Sources Included in the Narrative Synthesis

AUTHOR (YEAR)	STUDY DESIGN	POPULATION/DATA SOURCE	DIGITAL TECHNOLOGY	EQUITY RELEVANCE
Cui (2025)	Conceptual framework	Digital health data streams across populations	Digital health algorithms; distributional shift and representation bias mitigation	Links representation bias and distributional shift to generalizability and equity risks.
Chen et al. (2023)	Perspective review	AI/ machine learning (ML) applications across medical domains	Clinical AI and ML algorithms	Summarizes fairness trade-offs and bias sources across AI development stages.
Obermeyer et al. (2019)	Retrospective analysis	Large US health system commercial dataset	Commercial cost-based risk stratification algorithm	Demonstrates racial bias from a cost-based proxy for health need.
Benjamin (2019)	Critical commentary	Sociological analysis (no primary data)	Clinical risk algorithms	Frames algorithmic quantification as a mechanism that may legitimize structural discrimination.
Goodman et al. (2018)	Commentary and conceptual analysis	Healthcare utilization data	ML health prediction models	Explains why healthcare cost may reflect access rather than morbidity.
Paulus and Kent (2020)	Methodological analysis	Clinical prediction model literature	Algorithmic clinical prediction tools	Shows how imperfect proxy outcomes can generate unequal predictions.
Siddique et al. (2024)	Systematic review (63 studies)	Diverse health system algorithms	Healthcare algorithms broadly	Finds that algorithms may reduce, sustain, or worsen inequities depending on context.
Agarwal et al. (2023)	Framework development	ML pipeline literature	ML algorithms across health contexts	Proposes pipeline-level bias awareness and mitigation across development stages.
Embi (2021)	Viewpoint and framework proposal	Clinical AI deployment contexts	AI-driven healthcare systems	Proposes algorithmic vigilance for post-deployment monitoring of safety and equity.
Mhasawade et al. (2021)	Analytical review	Public and population health ML studies	Population health ML algorithms	Supports integrating social determinants into fairness-aware population health modeling.
Badr et al. (2024)	Scoping review (41 studies)	General population; studies examining digital health technology use and inequalities across sociodemographic groups	Digital health technologies broadly	Identifies sociodemographic patterns in digital health adoption and access barriers.
Zhang et al. (2023)	Population-level analysis	NHS administrative data	NHS digital health services	Shows area deprivation independently predicts lower digital service use.
Lee et al. (2023)	Scoping review (40 articles)	Digital phenotyping studies across settings	Smartphone-based phenotyping (mobility, sleep, communication frequency)	Reviews smartphone-derived features and their context-dependent interpretation.
Birk and Samuel (2020)	Sociological analysis	Digital phenotyping literature	Smartphone-based digital phenotyping	Critiques assumptions of normality, use of digital data as proxies for social life, and risk of reifying mental health problems.
Buda et al. (2022)	Empirical analysis	Commercial digital well-being platform users	Commercial well-being app algorithms	Applied a seven-step fairness framework to a commercial digital well-being service and pilot-tested subgroup fairness assessment among 610 users.
Zhen et al. (2024)	Cross-sectional survey analysis using CHARLS 2020 data	Nationally representative older adult cohort	Digital technology broadly (internet, smartphones)	Finds digital technology use associated with lower health inequality among older adults in China.
Cahan et al. (2019)	Perspective and framework	Big data health applications	AI and ML in precision medicine	Emphasizes dataset assessment before algorithm development.
Chin et al. (2023)	Consensus framework (multistakeholder)	Healthcare algorithm governance	Clinical algorithms broadly	Recommends transparent, participatory, and accountable algorithm governance.
Alderman et al. (2025)	Delphi consensus (350+ stakeholders)	Health AI dataset developers, regulators, and end users	Health AI datasets and algorithmic systems broadly	Provides consensus recommendations for transparent and diverse health AI datasets.
Busch et al. (2024)	Legal and regulatory analysis	EU member states; AI-enabled medical device sector	High-risk AI in healthcare; EU regulatory framework	Analyzes EU AI Act requirements and remaining equity-reporting gaps.
Colacci et al. (2025)	Scoping review (91 studies)	Clinical ML models across health systems	Clinical ML algorithms broadly	Documents high prevalence of sociodemographic bias and limited intersectional assessment.
Mackin et al. (2025)	Extended umbrella review	Healthcare binary classification models; electronic medical records	Post-processing bias mitigation tools for off-the-shelf ML models	Reviews post-processing mitigation tools and reports heterogeneous effectiveness across methods, metrics, and studies.

Note: Table 1 summarizes the 22 primary sources included in the narrative synthesis. Detailed thematic interpretation is provided in the Results and Discussion sections. Additional cited sources support background, methods, regulatory context, implementation implications, and comparison with adjacent frameworks but were not counted as primary synthesis sources. ML = machine learning; NHS = National Health Service; SES = socioeconomic status.

populations served. Zhen et al.^[35] provide evidence from China that digital technology use was associated with lower health inequality among older adults, though generalizability beyond that context remains uncertain. In LMICs, Sylla et al.^[36] highlight access, affordability, and infrastructure as important considerations for equitable digital health expansion. Osonuga et al.^[37] discuss AI as a potential catalyst for health equity in primary care while emphasizing the need to address digital divide and algorithmic equity risks. Structural barriers at the infrastructure, literacy, trust, and interface design levels translate into unequal engagement with digital tools, reducing data representativeness and limiting the reach of algorithmically mediated care.^[3,11,12,34] Together, these sources indicate that digital access barriers can function as deployment-stage mechanisms through which upstream social inequity becomes downstream algorithmic inequity.

Data Foundations and Structural Bias:

Cahan et al. argue that algorithmic development without critical dataset assessment may produce training data overrepresenting insured, engage patients and not reflecting true disease burden in underserved populations; training data representativeness should therefore be assessed before model development.^[38] Goodman et al.^[6] and Paulus and Kent^[7] explain how proxy outcomes such as healthcare costs can encode historical underinvestment rather than true clinical need. Celi et al.^[39] documented disparities in the geography, clinical specialty, authorship, and data-source composition of clinical AI studies, highlighting risks to generalizability and global health equity. Chen et al.^[2] further describe how data acquisition, dataset shift, and labelling variability can contribute to algorithmic unfairness. The STANDING Together consensus statement brought together 29 recommendations for improving algorithmic transparency and equity in health AI datasets, developed through a Delphi process involving 194 participants from 25 countries and input from more than 350 stakeholders across 58 countries.^[40] Together, these sources indicate that structural inequity can enter the digital health pipeline before model development begins through selection, measurement, annotation, and proxy outcome choices.

Lifecycle-Based Approaches and Frameworks for Accountability:

Algorithm governance should consider equity throughout the AI lifecycle. Chin et al.^[41] defined multi-stakeholder governance principles including transparency, community participation in fairness outcomes, and prospective accountability. Alderman et

al.^[40] operationalized these through STANDING Together. Embi^[8] proposed algorithmic vigilance, a post-deployment equity monitoring approach analogous to pharmacovigilance, noting that certain algorithm-related harms are only detectable after deployment. Davis et al.^[42] provided empirical support, demonstrating fairness drift in a national Veterans Affairs population over 11 years: fairness could not be assured through one-time developmental assessment, and model updating sometimes exacerbated rather than reduced fairness gaps. Agarwal et al.^[5] argued that bias mitigation is more likely to succeed when integrated throughout development rather than applied post-hoc. The EU AI Act has been described as the first comprehensive legal framework specifically focused on AI and is highly relevant to healthcare; however, in the authors' interpretation, it does not explicitly make demographic subgroup performance reporting or health-equity impact assessment stand-alone conditions for market authorization.^[21] Ferryman^[43] identified a parallel gap in the US: the FDA's AI/ML regulatory framework does not treat health disparity assessment as a stand-alone condition of good machine learning practice. Together, these sources suggest that governance mechanisms can either interrupt or permit the persistence of inequities introduced earlier in the pipeline.

Discussion

The synthesized evidence suggests that many digital health inequities reflect reproduction, amplification, or reconfiguration of pre-existing structural disadvantages rather than purely novel technical problems.^[1,2,6,20] The critical contribution of this review lies not in cataloguing individual harms alone but in arguing for their systemic interconnection and cumulative relevance.

The Structural Compounding Model of Digital Health Inequity:

The six thematic domains share a common structural logic: each represents a stage at which inequity can enter, persist, and amplify. The Structural Compounding Model formalizes this into a four-stage pipeline (Data, Model, Deployment, Governance), making explicit how structural disadvantage accrues across stages. Each stage carries inequities consistent with, and potentially compounding upon, those of prior stages; inter-level interaction strength will vary by context, technology, and population.

Levels:

[1] Data Level:

Sociodemographic characteristics and institutional

Table 2:
The Structural Compounding Model of Digital Health Inequity

Level 1: DATA	Level 2: MODEL	Level 3: DEPLOYMENT	Level 4: GOVERNANCE
<p>Unrepresentative training data may limit model generalizability and contribute to inequitable performance across populations.^[2,38,39]</p> <p>Cost-based proxies may reproduce historical inequities in healthcare access.^[1]</p> <p>Digital phenotyping features derived from mobility, sleep, activity, location, communication, and in-phone activity may be misinterpreted if social and environmental context is not considered.^[29,30]</p> <p>Data acquisition, labelling variability, dataset composition, and curation practices may contribute to inequities within training datasets.^[2,39]</p>	<p>Fairness criteria may not align across stakeholders or clinical objectives.^[2]</p> <p>Optimization for overall model performance may favour majority populations when subgroup performance is not explicitly evaluated.^[2]</p> <p>Distributional shift may reduce model performance and generalizability for underrepresented populations.^[24]</p> <p>Bias may be introduced or reinforced across different steps of the machine-learning pipeline.^[39]</p>	<p>Infrastructure, literacy, trust, and usability barriers may reduce the effectiveness and adoption of digital health technologies.^[11]</p> <p>Inequities in device ownership and internet access may limit access to digital health technologies.^[34]</p> <p>Digital health interface design may not be culturally or contextually compatible with all communities.^[45]</p> <p>Historical medical mistreatment, inequity, and mistrust may reduce digital health technology use among marginalized communities.^[46]</p> <p>Meeting technical fairness criteria alone does not guarantee equitable health outcomes.^[2]</p>	<p>Post-deployment monitoring may help assess the effectiveness, safety, and equity of health AI systems.^[8]</p> <p>Reporting subgroup performance and dataset composition may improve transparency about model behaviour and limitations across diverse groups.^[2,40,41]</p> <p>Community involvement and culturally responsive design may improve the acceptability and relevance of digital health solutions across population groups.^[45]</p> <p>Post-deployment monitoring should be paired with proactive equity assessment to identify inequities that emerge during deployment.^[8]</p>

structures in healthcare access shape the data available, how outcomes are classified, and which experiences become part of training data. People with reduced healthcare access may generate fewer data points and less representative data. Differences in healthcare cost may act as proxies for differences in access rather than actual differences in healthcare need.^[1,6,7] Digital phenotyping features may incorporate socioeconomic circumstances alongside clinical status, creating risk of misclassification if not validated across contexts.^[29-31] Bias present at this stage may persist throughout subsequent stages.^[5,38]

[2] Model Level:

Unequal representation in training data may contribute to unequal algorithm performance among underrepresented population segments. Multiple fairness criteria cannot always be optimized simultaneously, and such trade-offs may remain implicit or favour overall performance in majority populations.^[2] Distributional shift may reduce model performance for people who differ from the training data distribution.^[24]

[3] Deployment Level:

Even well-developed algorithms may have disparate effects in vulnerable communities where structural barriers such as limited digital literacy, insufficient device access, and poor internet connectivity can restrict equitable access to algorithmic services.^[11,12] Deployment is an inherently intersectional stage because factors such as race, social status, geographical location, age, and language can affect access and outcomes in digital health systems.^[3] Crawford et al.^[44] illustrate the importance of

intersectional and community-informed approaches in digital health intervention development among underserved women, highlighting the need to address gendered, social, and economic disparities concurrently in implementation.

[4] Governance Level:

At the Governance Level, absent or inadequate equity requirements in regulatory frameworks may allow inequities generated at prior stages to persist undetected and uncorrected.^[21,43] Reducing algorithmic bias without addressing the data on which models are trained may improve measured model performance without promoting population health equity.^[5] Critically, the Structural Compounding Model does not propose a single root cause of digital health inequity; rather, it holds that inequity can enter and amplify at any of the four levels, and that the specific combination of contributing factors varies by context, technology, and population.

The Application of the Structural Compounding Model:

Obermeyer et al.^[1] provide an illustrative case that can be interpreted, with appropriate caution, through all four levels of the proposed model: cost-based proxy targets encoded access inequity at the Data Level; Black patients were systematically under identified at the Model Level; use of biased outputs in care-management referral workflows could reduce equitable access to additional care at the Deployment Level; and the absence of mandatory subgroup reporting meant no governance mechanism required disclosure of the disparity at the Governance Level.

Heterogeneity of Algorithmic Influence on Equity:

The direction and magnitude of algorithmic influence on health equity cannot be assumed in advance. Siddique et al.^[9] found that healthcare algorithms could decrease, preserve, or exacerbate inequities depending on design considerations, outcome metrics, and contextual factors. This heterogeneity suggests that inequitable algorithmic effects are not inevitable but are linked to decisions made by developers, implementers, and users. Some evidence suggests that under appropriate structural conditions, digital technologies may help decrease, rather than increase, health inequities.^[35]

Digital Phenotyping as an Under-Studied Equity Issue:

Equity-specific evidence in digital phenotyping and digital mental health remains comparatively limited. Existing work includes scoping, sociological, and pilot fairness analyses, but broader validation across diverse populations and contexts remains insufficient; rapid digital mental health expansion without equity impact assessments represents a governance gap that current EU AI Act provisions may not adequately address.^[21,29-31]

How This Review Overcomes Gaps in Current Literature:

This review addresses three structural limitations of the existing literature. First, prior work is fragmented across ethics appraisal, AI lifecycle assessment, digital health implementation, dataset transparency, and algorithmic governance. Existing frameworks provide important guidance, including Suresh and Gutttag's taxonomy of machine-learning harms^[13], Char et al.'s ethics appraisal framework^[14], HEAAL's applied lifecycle equity assessment framework^[15], DH-EquIR's implementation research model^[16], AI/ML total product lifecycle equity considerations^[17], the evidence- and consensus-based Digital Healthcare Equity Framework^[18], and recent synthesis of telehealth and digital health equity frameworks.^[19] However, these frameworks generally emphasize assessment or implementation rather than theorizing structural inequity as a sequential compounding process. The Structural Compounding Model contributes a theory-generating public health framework that links upstream social disadvantage, data generation, model design, deployment conditions, and governance accountability within one sequential model. Second, single-level interventions may be insufficient because inequity roots span multiple levels simultaneously; governance frameworks by Chin et al.^[41], Alderman et al.^[40], and Embj^[8] may not compensate for upstream data and model deficiencies. Third, only 1.1% of bias assessments incorporated

intersectional analysis;^[20] the Structural Compounding Model addresses this by treating deployment as inherently intersectional, with race, socioeconomic status, geography, age, and language as co-determining access conditions. Inadequate governance may allow inequities to persist and may reduce incentives for proactive equity assessment when regulatory requirements do not explicitly require such assessment.^[21,43]

The model is not intended to replace operational frameworks such as HEAAL, DH-EquIR, the Digital Healthcare Equity Framework, or total product lifecycle approaches. Rather, it provides an explanatory structure for understanding why isolated technical, regulatory, or implementation interventions may be insufficient when inequity originates across multiple linked stages. Its main contribution is theoretical integration: it connects upstream social disadvantage, data generation, model design, deployment conditions, and governance accountability within one sequential model.

Limitations

Methodological Limitations of the Included Articles:

Consistent with patterns documented in algorithmic bias reviews, many included empirical studies use retrospective designs conducted at a single site, restricting generalizability beyond the health system studied; only a minority include multi-site or nationally representative data.^[9,20] Fairness measures focus predominantly on individual demographic characteristics rather than intersectional combinations of race, class, geography, gender, and age, limiting the depth of equity analysis possible.

Limitations of the Present Review:

This review is subject to limitations inherent to a single-reviewer narrative design, including susceptibility to selection and thematic coding biases. Future reviews should employ dual screeners with reliability procedures. The review did not include inter-coder reliability assessment, formal risk-of-bias scoring, protocol registration, or exhaustive grey literature searching. Heterogeneity among included sources prevented meta-analysis. The dominance of high-income country literature limits global applicability. Google Scholar was used as a supplementary citation-checking and relevance-scanning source but yielded no additional eligible source after deduplication and relevance assessment. Five sources identified through targeted citation searches were retained within the final synthesis because they were directly relevant to framework development. The Structural Compounding

Model should be interpreted as a theory-generating framework that requires empirical validation across specific digital health technologies, populations, and health-system contexts.

Implications

The following implications are derived directly from the synthesized findings, organized by stakeholder group.

For researchers: Only 1.1% of bias assessments include intersectional analysis;^[20] study designs should incorporate a priori intersectional subgroup analysis. Proxy outcomes reflecting access rather than need^[1,6,7] require critical evaluation before model training.

For practitioners and health systems: Evidence that algorithms can mitigate, maintain, or exacerbate inequities depending on design and deployment context^[9] supports the use of equity-focused impact assessment before implementation and continuous post-deployment monitoring after implementation.^[8]

For policy makers: In the authors' interpretation, the EU AI Act does not explicitly make pre-market equity assessment or demographic subgroup reporting stand-alone conditions for market authorization, representing a potential legislative gap.^[21] Regulatory frameworks could strengthen equity oversight by requiring pre- and post-market equity evaluation as a condition of market authorization for high-risk AI systems.^[40,41] The consistent association between area-level deprivation and lower digital service uptake identifies infrastructure investment and digital literacy programs as a priority area consistently highlighted across the access literature.^[11,12,36]

For developers and platforms: Evidence and expert frameworks indicate that bias mitigation should be integrated across the development and deployment lifecycle rather than treated only as a post-hoc correction^[5]; accordingly, community-centred co-design and equity-sensitive performance metrics should be embedded from the outset of algorithm development.^[41,45,46]

Future Research Priorities

The following section identifies evidence gaps from this review that represent priorities for future empirical research. First, this review did not identify prospective, multi-centre studies with a priori intersectional subgroup analysis in this domain; such studies are needed.^[20] Second, binary pairwise demographic comparisons dominate the fairness literature; intersectional measures simultaneously accounting for race, class, geography, gender, and age require development and validation. Third, the equity

implications of digital phenotyping have been examined primarily in theoretical or small, non-representative samples^[29,30]; nationally representative prospective studies accounting for socioeconomic status, housing security, employment, and cultural context are required. Fourth, post-processing bias mitigation tool performance in lower-resource health settings has not been adequately studied^[47]; this remains an important gap given global adoption of off-the-shelf algorithms. Fifth, evidence on digital health equity in low- and middle-income countries remains limited despite rapid technology adoption^[36]; this geographic gap is an important research priority. Sixth, common standards for measuring and documenting equity outcomes across the digital health tool lifecycle remain absent^[40]; developing such standards would enable cumulative evidence synthesis.

Conclusion

Algorithmic bias and health-based digital inequities often reflect structural processes, including long-term underinvestment, biased data collection, proxy outcomes, and unequal access to digital technologies, rather than merely technical problems. The Structural Compounding Model of Digital Health Inequity illustrates how these elements may interact: inequities entering at any of the four pipeline levels (Data, Model, Deployment, and Governance) may not remain isolated but may compound in ways that sustain exclusion of groups already disadvantaged within health care. Some evidence suggests that under appropriate structural conditions, digital health interventions may reduce rather than widen health inequities, indicating that compounding disadvantage is shaped by design, implementation, infrastructure, and governance choices. Equitable digital health outcomes are likely to require an integrated, multi-tier approach throughout the entire pipeline: fair data governance practices; transparent balancing of fairness trade-offs; digital access initiatives; and stronger equity assessment, subgroup demographic reporting, and post-implementation algorithmic oversight.

Declaration

Funding: No funding was received for this research.

Conflict of Interest: No conflicts of interest are declared.

AI Tool Disclosure: Artificial intelligence tools were used for grammar and clarity.

References

1. A Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage

- the health of populations. *Science*. 2019;366(6464):447-453. doi:10.1126/science.aax2342
2. Chen RJ, Wang JJ, Williamson DFK, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat Biomed Eng*. 2023;7(6):719-742. doi:10.1038/s41551-023-01056-8
 3. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. *NPJ Digit Med*. 2021;4(1):52. doi:10.1038/s41746-021-00413-8
 4. Marmot MG, Wilkinson RG, eds. *Social Determinants of Health*. 2nd ed. Oxford: Oxford University Press; 2006.
 5. Agarwal R, Bjarnadottir M, Rhue L, et al. Addressing algorithmic bias and the perpetuation of health inequities: an AI bias aware framework. *Health Policy Technol*. 2023;12(1):100702. doi:10.1016/j.hlpt.2022.100702
 6. Goodman SN, Goel S, Cullen MR. Machine learning, health disparities, and causal reasoning. *Ann Intern Med*. 2018;169(12):883-884. doi:10.7326/M18-3297
 7. Paulus JK, Kent DM. Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *NPJ Digit Med*. 2020;3:99. doi:10.1038/s41746-020-0304-9
 8. Embi PJ. Algorithmovigilance: advancing methods to analyze and monitor artificial intelligence-driven health care for effectiveness and equity. *JAMA Netw Open*. 2021;4(4):e214622. doi:10.1001/jamanetworkopen.2021.4622
 9. Siddique SM, Tipton K, Leas B, et al. The impact of health care algorithms on racial and ethnic disparities: a systematic review. *Ann Intern Med*. 2024;177(4):484-496. doi:10.7326/M23-2960
 10. Benjamin R. Assessing risk, automating racism: a health care algorithm reflects underlying racial bias in society. *Science*. 2019;366(6464):421-422. doi:10.1126/science.aaz3873
 11. Badr J, Motulsky A, Denis JL. Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations. *Health Policy*. 2024;146:105122. doi:10.1016/j.healthpol.2024.105122
 12. Zhang J, Gallifant J, Pierce RL, et al. Quantifying digital health inequality across a national healthcare system. *BMJ Health Care Inform*. 2023;30(1):e100809. doi:10.1136/bmjhci-2023-100809
 13. Suresh H, Gutttag JV. A framework for understanding sources of harm throughout the machine learning life cycle. In: *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21)*. New York: Association for Computing Machinery; 2021. doi:10.1145/3465416.3483305
 14. Char DS, Abramoff MD, Feudtner C. Identifying ethical considerations for machine learning healthcare applications. *Am J Bioeth*. 2020;20(11):7-17. doi:10.1080/15265161.2020.1819469
 15. Kim JY, Hasan A, Kellogg KC, et al. Development and preliminary testing of Health Equity Across the AI Lifecycle (HEAAL): a framework for healthcare delivery organizations to mitigate the risk of AI solutions worsening health inequities. *PLOS Digit Health*. 2024;3(5):e0000390. doi:10.1371/journal.pdig.0000390
 16. Groom LL, Schoenthaler AM, Mann DM, Brody AA. Construction of the Digital Health Equity-Focused Implementation Research Conceptual Model: bridging the divide between equity-focused digital health and implementation research. *PLOS Digit Health*. 2024;3(5):e0000509. doi:10.1371/journal.pdig.0000509
 17. Abramoff MD, Tarver ME, Loyo-Berrios N, et al. Considerations for addressing bias in artificial intelligence for health equity. *NPJ Digit Med*. 2023;6(1):170. doi:10.1038/s41746-023-00913-9
 18. Hatf E, Scholle SH, Buckley B, Weiner JP, Austin JM. Development of an evidence- and consensus-based Digital Healthcare Equity Framework. *JAMIA Open*. 2024;7(4):ooae136. doi:10.1093/jamiaopen/ooae136
 19. Wang SW, Killedear A, Von Huben A, et al. Evaluation of health equity frameworks in telehealth and digital health: a systematic review and narrative synthesis. *Front Public Health*. 2025;13:1690117. doi:10.3389/fpubh.2025.1690117
 20. Colacci M, Huang YQ, Postill G, et al. Sociodemographic bias in clinical machine learning models: a scoping review of algorithmic bias instances and mechanisms. *J Clin Epidemiol*. 2025;178:111606. doi:10.1016/j.jclinepi.2024.111606
 21. Busch F, Kather JN, Johner C, et al. Navigating the European Union Artificial Intelligence Act for healthcare. *NPJ Digit Med*. 2024;7(1):210. doi:10.1038/s41746-024-01213-6
 22. Baethge C, Goldbeck-Wood S, Mertens S. SANRA: a scale for the quality assessment of narrative review articles. *Res Integr Peer Rev*. 2019;4:5. doi:10.1186/s41073-019-0064-8
 23. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101. doi:10.1191/1478088706qp063oa
 24. Cui Y. Digital pathways connecting social and biological factors to health outcomes and equity. *NPJ Digit Med*. 2025;8:172. doi:10.1038/s41746-025-01564-8
 25. Mhasawade V, Zhao Y, Chunara R. Machine learning and algorithmic fairness in public and population health. *Nat Mach Intell*. 2021;3(8):659-666. doi:10.1038/s42256-021-00373-4
 26. Charpignon ML, Carrel A, Jiang Y, et al. Going beyond the means: exploring the role of bias from digital determinants of health in technologies. *PLOS Digit Health*. 2024;3(5):e0000390. doi:10.1371/journal.pdig.0000390

- Health. 2023;2(10):e0000244. doi:10.1371/journal.pdig.0000244
27. Otokiti AU, Shih HJ, Williams KS. Gender and racial bias unveiled: clinical artificial intelligence and machine learning algorithms are fanning the flames of inequity. *Oxf Open Digit Health*. 2025;3:0qaf027. doi:10.1093/oodh/oqaf027
28. Hussain SA, Bresnahan M, Zhuang J. The bias algorithm: how AI in healthcare exacerbates ethnic and racial disparities - a scoping review. *Ethn Health*. 2025;30(2):197-214. doi:10.1080/13557858.2024.2422848
29. Lee KG, Lee TC, Yefimova M, et al. Using digital phenotyping to understand health-related outcomes: a scoping review. *Int J Med Inform*. 2023;174:105061. doi:10.1016/j.ijmedinf.2023.105061
30. Birk RH, Samuel G. Can digital data diagnose mental health problems? A sociological exploration of digital phenotyping. *Sociol Health Illn*. 2020;42(8):1873-1887. doi:10.1111/1467-9566.13175
31. Buda T, Guerreiro J, Iglesias J, et al. Foundations for fairness in digital health apps. *Front Digit Health*. 2022;4:943514. doi:10.3389/fdgth.2022.943514
32. Rowe S, Newby D, Wykes T. Digital phenotyping: how it could change mental health care and why we should all keep up. *J Ment Health*. 2024;33(4):439-442. doi:10.1080/09638237.2024.2395537
33. Robinson A, Flom M, Forman-Hoffman VL, et al. Equity in digital mental health interventions in the United States: where to next? *J Med Internet Res*. 2024;26:e59939. doi:10.2196/59939
34. Yao R, Zhang W, Evans R, Cao G, Rui T, Shen L. Inequities in health care services caused by the adoption of digital health technologies: scoping review. *J Med Internet Res*. 2022;24(3):e34144. doi:10.2196/34144
35. Zhen Z, Tang D, Wang X, Feng Q. The impact of digital technology on health inequality: evidence from China. *BMC Health Serv Res*. 2024;24(1):1531. doi:10.1186/s12913-024-12022-8
36. Sylla B, Ismaila O, Diallo G. 25 years of digital health toward universal health coverage in low- and middle-income countries: rapid systematic review. *J Med Internet Res*. 2025;27:e59042. doi:10.2196/59042
37. Osonuga A, Osonuga AA, Fidelis SC, et al. Bridging the digital divide: artificial intelligence as a catalyst for health equity in primary care settings. *Int J Med Inform*. 2025;204:106051. doi:10.1016/j.ijmedinf.2025.106051
38. Cahan EM, Hernandez-Boussard T, Thadaney-Israni S, Rubin DL. Putting the data before the algorithm in big data: addressing personalized healthcare. *NPJ Digit Med*. 2019;2(1):78. doi:10.1038/s41746-019-0157-2
39. Celi LA, Cellini J, Charpignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities: a global review. *PLOS Digit Health*. 2022;1(3):e0000022. doi:10.1371/journal.pdig.0000022
40. Alderman JE, Palmer J, Laws E, et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. *Lancet Digit Health*. 2025;7(1):e64-e88. doi:10.1016/S2589-7500(24)00224-3
41. Chin MH, Afsar-Manesh N, Bierman AS, et al. Guiding principles to address the impact of algorithm bias on racial and ethnic disparities in health and health care. *JAMA Netw Open*. 2023;6(12):e2345050. doi:10.1001/jamanetworkopen.2023.45050
42. Davis SE, Dorn C, Park DJ, Matheny ME. Emerging algorithmic bias: fairness drift as the next dimension of model maintenance and sustainability. *J Am Med Inform Assoc*. 2025;32(5):845-854. doi:10.1093/jamia/ocaf039
43. Ferryman K. Addressing health disparities in the Food and Drug Administration's artificial intelligence and machine learning regulatory framework. *J Am Med Inform Assoc*. 2020;27(12):2016-2019. doi:10.1093/jamia/ocaa133
44. Crawford AD, Slavin R, Tabar M, et al. Methodological approaches in developing and implementing digital health interventions amongst underserved women. *Public Health Nurs*. 2024;41(6):1612-1621. doi:10.1111/phn.13410
45. Naderbagi A, Loblay V, Zahed IUM, et al. Cultural and contextual adaptation of digital health interventions: narrative review. *J Med Internet Res*. 2024;26:e55130. doi:10.2196/55130
46. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. *NPJ Digit Med*. 2022;5(1):119. doi:10.1038/s41746-022-00663-0
47. Mackin S, Major VJ, Chunara R. Post-processing methods for mitigating algorithmic bias in healthcare classification models: an extended umbrella review. *BMC Digit Health*. 2025;3:26. doi:10.1186/s44247-025-00166-4