https://doi.org/10.70844/jmhrp.2025.2.2.38

Journal of Medical Health Research and Psychiatry

Research Article

JMHRP-25-045

A Systematic Framework for Investigating Algorithmic Bias as a Social Determinant of Health in Low and Middle-Income Countries

Anthony Iyore Anaweokhai*

Department of Public Health, Anaweokhai Institute for Applied Global Health, Nigeria

Corresponding Author: Anthony Iyore Anaweokhai, Department of Public Health, Anaweokhai Institute for Applied Global Health, Nigeria, E-mail: anaweokhai.institute@gmail.com

Received date: 23 July, 2025, **Accepted date:** 01 September, 2025, **Published date:** 09 September, 2025

Citation: Anaweokhai AI (2025) Research Protocol: A Systematic Framework for Investigating Algorithmic Bias as a Social Determinant of Health in Low and Middle-Income Countries. J Med Health Psychiatry 2: 2.

Abstract

Background: The rapid adoption of Artificial Intelligence (AI) and Machine Learning (ML) technologies in healthcare systems across Low and Middle-Income Countries (LMICs) presents unprecedented opportunities for improving health outcomes while simultaneously introducing novel risks for perpetuating and amplifying health inequities. Despite growing concerns about algorithmic bias in healthcare delivery, systematic methodological approaches for investigating these phenomena in LMIC contexts remain underdeveloped. Existing research frameworks, predominantly designed for high-income country settings, inadequately address the unique socioeconomic, cultural and infrastructural challenges that characterize LMIC healthcare systems.

Objective: This research protocol presents a comprehensive, multi-phase methodological framework for systematically investigating algorithmic bias as a mechanism through which social determinants of health operate in LMIC healthcare contexts. The protocol aims to establish standardized approaches for cross-country comparative research while maintaining sensitivity to local contexts and resource constraints.

Methods: We propose a sequential explanatory mixed-methods research protocol implemented across three phases over 36 months. Phase 1 involves systematic evidence mapping through scoping reviews and policy landscape analysis across 15 LMICs. Phase 2 comprises primary data collection through multi-stakeholder interviews (n=125) and in-depth healthcare system case studies in five countries (Nigeria, India, Kenya, Brazil, Bangladesh). Phase 3 focuses on framework development, validation and refinement through pilot implementation in Ghana and Vietnam.

Expected Outcomes: This protocol will yield a validated research framework for investigating algorithmic bias in LMIC healthcare systems, standardized measurement tools and indicators, evidence-based policy recommendations and capacity building guidelines. The methodology is designed to be culturally appropriate, ethically sound and implementable within resource-constrained settings.

Significance: By providing the first comprehensive research protocol for investigating algorithmic bias as a social determinant of health in LMICs, this work establishes the methodological foundation for future empirical studies, evidence-based policy development and international collaborative research efforts aimed at ensuring health equity in the digital health era.

Keywords: Algorithmic bias, Social determinants of health, Digital health, Health equity, Low and middle-income countries, Research methodology, Artificial intelligence

Introduction

The digital health transformation in LMICs

The global health landscape is experiencing an unprecedented digital transformation, with artificial intelligence and machine learning technologies increasingly integrated into healthcare delivery systems worldwide [1]. This technological revolution is particularly pronounced in low- and middle-income countries, where digital health innovations offer opportunities to leapfrog traditional healthcare infrastructure limitations and address persistent health system challenges [2]. Mobile health platforms, AI-powered diagnostic tools and algorithmic clinical decision support systems are being rapidly deployed across diverse LMIC contexts, from rural health centers in sub-Saharan Africa to urban hospitals in Southeast Asia [3].

However, this rapid technological adoption occurs within complex socioeconomic environments characterized by significant health inequities, limited regulatory frameworks and varying degrees of digital literacy among both healthcare providers and patients [4]. Unlike high-income countries where AI implementation typically occurs within well-established healthcare systems with robust oversight mechanisms, LMICs often lack the institutional capacity to

https://doi.org/10.70844/jmhrp.2025.2.2.38

Journal of Medical Health Research and Psychiatry

adequately evaluate, monitor and mitigate potential algorithmic biases embedded within these technologies [5]. This creates a paradoxical situation where the populations most vulnerable to health inequities may be simultaneously most exposed to algorithmic systems that could perpetuate or amplify existing disparities.

Current evidence regarding algorithmic bias in healthcare derives predominantly from high-income country contexts, with limited empirical data available from LMIC settings [6]. Existing research methodologies developed for studying algorithmic bias in well-resourced healthcare systems may be inadequately suited for LMIC contexts, where data availability, regulatory frameworks and healthcare delivery models differ substantially from those in high-income countries [7]. This methodological gap represents a critical barrier to understanding how algorithmic bias manifests in LMIC healthcare systems and developing appropriate mitigation strategies.

The absence of systematic research approaches specifically designed for investigating algorithmic bias in LMIC contexts has profound implications for global health equity. Without robust empirical evidence and standardized methodological frameworks, policymakers and healthcare leaders in LMICs lack the evidence base necessary to make informed decisions about AI implementation, regulation and governance [8]. This knowledge gap is particularly concerning given the potential for algorithmic bias to operate as a mechanism through which traditional social determinants of health are amplified, creating novel pathways for health inequity in the digital age.

Algorithmic bias and health equity: Current understanding

Algorithmic bias in healthcare encompasses systematic and unfair discrimination embedded within computational systems that affects clinical decision-making, resource allocation and health service delivery [9]. These biases can manifest through various mechanisms, including biased training data, inappropriate algorithmic design choices and inadequate consideration of population diversity during system development and deployment [5]. Recent high-profile cases, such as the identification of racial bias in commercial healthcare algorithms used for patient risk stratification, have highlighted the potential for AI systems to perpetuate existing healthcare disparities while appearing objective and neutral [9].

The intersection between algorithmic bias and social determinants of health represents a critical area of investigation that remains theoretically underdeveloped and empirically understudied [10]. Social determinants of health, as conceptualized by the World Health Organization, encompass the conditions in which people are born, grow, live, work and age, including the broader set of forces and systems that shape daily life conditions [11]. These determinants operate through complex pathways to influence health outcomes, with structural factors such as socioeconomic stratification, governance systems and cultural norms creating differential exposures and vulnerabilities across population groups.

Algorithmic bias in healthcare may function as a novel mechanism through which traditional social determinants of health operate in digitalized healthcare systems [12]. For example, algorithmic systems trained on historical healthcare data may embed patterns of differential care access and quality that reflect underlying socioeconomic inequities. When deployed in clinical settings, these systems may systematically underestimate disease risk or treatment needs among marginalized populations, effectively institutionalizing discriminatory practices within seemingly objective technological tools [13].

In LMIC contexts, the potential for algorithmic bias to amplify existing health inequities is particularly concerning given the intersection of digital divides with traditional forms of social stratification [4]. Limited internet connectivity, low digital literacy rates and insufficient representation of LMIC populations in AI training datasets create conditions where algorithmic systems may systematically underperform for the populations they are intended to serve [14]. Furthermore, the rapid scaling of AI technologies in resource-constrained environments may occur without adequate consideration of local contexts, cultural factors and community needs, potentially exacerbating rather than addressing health inequities.

Despite the theoretical importance of understanding algorithmic bias as a social determinant of health in LMIC contexts, empirical evidence remains limited and methodologically fragmented. Existing studies typically focus on technical aspects of bias detection and mitigation in high-income country settings, with limited attention to the social, cultural and economic factors that shape bias manifestation and impact in diverse global contexts [6]. This evidence gap represents a critical barrier to developing effective policies and practices for ensuring health equity in the digital health era.

Methodological challenges in LMIC research contexts

Investigating algorithmic bias in LMIC healthcare systems presents unique methodological challenges that require specialized research approaches [7]. Data availability and quality represent fundamental constraints, as many LMIC healthcare systems lack comprehensive electronic health records, standardized data collection protocols and robust information management systems [15]. This data scarcity limits opportunities for retrospective analysis of algorithmic performance and bias detection, while also constraining the development of locally appropriate AI systems.

Regulatory and ethical considerations add additional complexity to research implementation in LMIC settings. Many countries lack established frameworks for AI governance in healthcare, creating uncertainty regarding appropriate ethical review processes, data protection requirements and consent procedures for algorithmic bias research [4]. Cultural sensitivity and community engagement represent critical considerations that may not be adequately addressed by standard research protocols developed for high-income country contexts.

Resource constraints, including limited research infrastructure, technical expertise and funding availability, further complicate the implementation of comprehensive algorithmic bias studies in LMIC settings [8]. Traditional research approaches may be prohibitively expensive or technically infeasible in resource-constrained environments, necessitating the development of pragmatic methodological frameworks that can generate robust evidence while remaining implementable within existing capacity limitations.

These methodological challenges underscore the critical need for specialized research protocols designed specifically for investigating algorithmic bias in LMIC healthcare contexts. Such protocols must balance scientific rigor with practical feasibility, cultural appropriateness with methodological standardization and local relevance with international comparability

https://doi.org/10.70844/jmhrp.2025.2.2.38

Journal of Medical Health Research and Psychiatry

Research Protocol Objectives

Primary objective

The primary objective of this research protocol is to develop a comprehensive, replicable methodological framework for systematically investigating algorithmic bias as a mechanism affecting health equity in LMIC healthcare systems. This framework will provide standardized approaches for identifying, measuring and analyzing algorithmic bias while maintaining sensitivity to diverse cultural, economic and healthcare system contexts across LMICs.

Secondary objectives

This research protocol encompasses several secondary objectives designed to address critical knowledge gaps and methodological limitations in current algorithmic bias research. First, the protocol aims to establish standardized methodological approaches that enable meaningful cross-country comparisons while accounting for contextual variations in healthcare delivery models, regulatory frameworks and socioeconomic conditions [8]. These standardized approaches will facilitate the development of comparative analyses that can identify common patterns of bias manifestation across diverse LMIC contexts while highlighting context-specific factors that require tailored interventions.

Second, the protocol seeks to create comprehensive frameworks for stakeholder engagement that recognize the complex multi-stakeholder environment characterizing LMIC healthcare systems [15]. These frameworks will address the involvement of diverse actors including government health officials, healthcare providers, technology developers, civil society organizations and community representatives, ensuring that research processes are participatory, culturally appropriate and responsive to local priorities and concerns.

Third, the protocol aims to develop culturally appropriate and linguistically accessible data collection instruments that can generate valid and reliable data across diverse LMIC populations [4]. These instruments will undergo rigorous cultural adaptation and validation processes to ensure their appropriateness for use across different cultural contexts, languages and healthcare settings while maintaining methodological consistency for comparative analysis.

Fourth, the protocol seeks to design implementation strategies that account for resource constraints and capacity limitations commonly encountered in LMIC research environments [7]. These strategies will prioritize cost-effective approaches, leverage existing infrastructure and partnerships and build local research capacity to ensure sustainable implementation and knowledge transfer.

Finally, the protocol aims to establish comprehensive ethical guidelines specifically tailored to algorithmic bias research involving vulnerable populations in LMIC settings. These guidelines will address issues of data sovereignty, benefit-sharing, community consent and protection of vulnerable groups while ensuring compliance with international ethical standards and local regulatory requirements.

Expected research questions to be addressed

This research protocol is designed to generate evidence addressing several critical research questions regarding algorithmic bias in LMIC healthcare contexts. Primary research questions include: How does algorithmic bias manifest in LMIC healthcare delivery systems and what are the specific mechanisms through which bias affects health outcomes among different population groups? What role do socioeconomic, cultural and linguistic factors play in moderating the effects of algorithmic bias on health equity? How do variations in healthcare system structure, governance and resource availability across LMICs influence patterns of algorithmic bias manifestation and impact?

Secondary research questions encompass policy and intervention-focused inquiries: What policy interventions and governance mechanisms can effectively mitigate identified biases while promoting beneficial AI adoption in LMIC healthcare systems? How can stakeholder engagement processes be optimized to ensure community participation in AI governance and bias mitigation efforts? What capacity building approaches are most effective for developing local expertise in algorithmic bias identification and mitigation? How can international cooperation and knowledge sharing be structured to support equitable AI development and deployment across diverse LMIC contexts?

Theoretical Framework and Conceptual Model

Social determinants of health: Foundation and evolution

The social determinants of health framework, as conceptualized by the World Health Organization Commission on Social Determinants of Health, provides the theoretical foundation for understanding how algorithmic bias may function as a mechanism affecting health equity [11]. This framework identifies structural determinants including socioeconomic stratification, governance systems, cultural norms and social policies as fundamental drivers of health inequities through their influence on intermediary determinants such as healthcare access, quality and utilization patterns.

Recent theoretical developments have expanded this framework to incorporate digital and technological factors as emerging social determinants of health [10]. Digital divides, algorithmic decision-making systems and technology-mediated healthcare delivery represent novel pathways through which traditional social determinants may operate in contemporary healthcare systems [12]. However, empirical investigation of these theoretical propositions remains limited, particularly in LMIC contexts where digital health implementation occurs within complex socioeconomic environments.

The integration of algorithmic systems into healthcare delivery creates new intermediary pathways between structural determinants and health outcomes. Algorithmic bias may function as a mechanism through which existing social stratification systems are embedded within technological systems, potentially amplifying traditional forms of discrimination while appearing objective and neutral [13]. This integration requires systematic investigation to understand how algorithmic bias interacts with traditional social determinants to affect health equity outcomes.

Proposed investigative framework

This research protocol proposes a novel theoretical framework conceptualizing algorithmic bias as an emergent social determinant of health that operates through both direct and indirect pathways to influence health outcomes in LMIC contexts. The core hypothesis underlying this framework posits that algorithmic bias functions as a mechanism through which traditional social determinants of health are amplified, modified, or transformed within digitalized healthcare systems.

The proposed conceptual model identifies three primary pathways through which algorithmic bias may affect health equity. The first pathway involves direct discrimination, where algorithmic systems systematically provide differential treatment recommendations, risk assessments, or resource allocations based on protected characteristics or their proxies [9]. This pathway operates through explicit or implicit consideration of demographic factors that correlate with social position and health outcomes.

The second pathway encompasses indirect discrimination, where algorithmic systems perpetuate existing healthcare disparities by embedding historical patterns of differential care access and quality within automated decision-making processes [5]. This pathway may be particularly relevant in LMIC contexts where historical healthcare data reflects significant inequities in service provision across different population groups.

The third pathway involves systemic exclusion, where certain population groups are systematically underrepresented in algorithmic system development, validation and deployment processes, resulting in systems that inadequately serve their needs [14]. This pathway may be especially pronounced in LMIC contexts where AI development often occurs in high-income countries with limited representation of LMIC populations in training datasets.

Key variables for empirical investigation include algorithmic system characteristics (type, complexity, deployment context), population characteristics (demographic, socioeconomic, cultural factors), healthcare system factors (structure, governance, resource availability) and outcome measures (clinical outcomes, care quality, access equity, patient satisfaction). Measurable indicators encompass both quantitative metrics (diagnostic accuracy disparities, treatment recommendation variations, resource allocation patterns) and qualitative assessments (stakeholder perceptions, community experiences, policy implementation challenges).

Context-specific considerations for LMICs

The application of this theoretical framework to LMIC contexts requires careful consideration of contextual factors that may influence how algorithmic bias manifests and operates within different healthcare systems. Healthcare system variations across LMICs include differences in public-private service provision, health financing mechanisms, regulatory frameworks and technology adoption patterns [15]. These variations may create different risk profiles and intervention opportunities for addressing algorithmic bias.

Cultural and linguistic diversity represents another critical contextual consideration, as algorithmic systems developed primarily in English or for Western populations may inadequately serve diverse LMIC communities with different languages, cultural practices and health beliefs [4]. Traditional healing systems, community health practices and cultural concepts of illness and wellness may not be adequately represented in AI systems designed according to biomedical paradigms.

Resource availability and technological infrastructure vary significantly across and within LMICs, creating differential capacity for implementing, monitoring and governing AI systems in healthcare [3]. Limited internet connectivity, inadequate technical expertise and constrained financial resources may affect both the deployment of AI systems and the implementation of bias mitigation strategies.

Methodology: Multi-Phase Research Protocol

Overall study design

This research protocol employs a sequential explanatory mixed-methods design implemented across three interconnected phases over a 36-month period. The methodology combines systematic evidence synthesis, primary data collection and framework development to generate comprehensive understanding of algorithmic bias manifestation and impact in LMIC healthcare contexts. The multi-phase approach enables iterative refinement of research questions, methodological approaches and theoretical frameworks based on emerging evidence and stakeholder feedback.

The research design incorporates a multi-country comparative framework encompassing diverse LMIC contexts representing different geographical regions, economic development levels, healthcare system structures and digital health implementation approaches. This comparative approach enables identification of common patterns while highlighting context-specific factors that influence algorithmic bias manifestation and impact.

Quality assurance mechanisms throughout the research process include regular methodology review meetings, external expert consultation, stakeholder validation workshops and systematic documentation of methodological decisions and adaptations. Data management procedures comply with international standards for research data governance, including provisions for data security, participant confidentiality and long-term data preservation.

Phase 1: Systematic evidence mapping

Scoping review protocol

The systematic evidence mapping component begins with a comprehensive scoping review designed to map existing knowledge regarding algorithmic bias in LMIC healthcare contexts and identify critical evidence gaps requiring primary investigation. The scoping review follows established methodological guidelines including the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) framework [16].

The literature search strategy encompasses multiple electronic databases including PubMed, Embase, Scopus, IEEE Xplore, Web of Science and regional databases such as African Index Medicus, LILACS and Index Medicus for Southeast Asia. Additional grey literature sources include government reports, international organization publications, conference proceedings and organizational websites. Search terms combine concepts related to algorithmic bias (artificial intelligence, machine learning, algorithm, bias, fairness, discrimination), healthcare applications (health, healthcare, medical, clinical, diagnosis, treatment) and LMIC contexts (developing country, low-income, middle-income, resource-limited, specific country names).

The temporal scope covers publications from 2015 to present, reflecting the emergence of AI applications in healthcare and growing recognition of algorithmic bias concerns. Language restrictions are limited to English, Spanish, French and Portuguese to maximize

https://doi.org/10.70844/jmhrp.2025.2.2.38

Journal of Medical Health Research and Psychiatry

coverage while maintaining feasibility within available translation resources.

Inclusion criteria encompass peer-reviewed articles, conference papers, government reports and organizational publications that address algorithmic bias, fairness, or discrimination in healthcare AI applications within LMIC settings. Studies must involve actual AI/ML applications in healthcare delivery, health system management, or public health practice. Exclusion criteria include purely theoretical papers without empirical content, studies from high-income countries without explicit LMIC relevance and publications focused solely on technical AI development without healthcare applications.

Data extraction utilizes a standardized form capturing study characteristics (design, setting, sample), AI application details (type, purpose, implementation context), bias assessment methods, findings and policy implications. Two independent reviewers conduct screening and data extraction with disagreements resolved through consensus discussion or third reviewer consultation.

Policy landscape analysis

The policy landscape analysis provides systematic examination of existing AI governance frameworks, regulatory approaches and policy initiatives addressing algorithmic bias in healthcare across selected LMIC contexts. This component employs document analysis methodology to assess policy content, implementation approaches and governance mechanisms.

Country selection employs purposive sampling to ensure representation across geographical regions (sub-Saharan Africa, South Asia, Southeast Asia, Latin America, Middle East), economic development levels (low-income, lower-middle-income, upper-middle-income) and healthcare system types (primarily public, mixed public-private, insurance-based). Selected countries include Nigeria, Ghana, Kenya, India, Bangladesh, Vietnam, Indonesia, Brazil, Mexico, Colombia, Egypt and Jordan.

Data sources encompass government websites, ministry of health publications, regulatory agency documents, national AI strategies, digital health policies and reports from international organizations such as WHO, World Bank and regional development banks. Additional sources include academic publications analyzing policy developments, civil society organization reports and industry association documents.

The analytical framework employs systematic content analysis using a standardized coding scheme addressing policy scope, governance mechanisms, stakeholder involvement, implementation timelines and enforcement provisions. Specific attention focuses on provisions addressing algorithmic bias, health equity, vulnerable population protection and international cooperation. Comparative analysis identifies policy patterns, best practices and implementation gaps across different country contexts.

Phase 2: Primary data collection

Multi-stakeholder interview protocol

The primary data collection phase employs semi-structured interviews with diverse stakeholders across multiple LMIC contexts to understand perceptions, experiences and practices related to algorithmic bias in healthcare. The interview methodology is designed to generate rich qualitative data while maintaining consistency across different cultural and linguistic contexts.

Participant recruitment employs purposive sampling with maximum variation to ensure representation across stakeholder groups, geographical locations and organizational types. Target participant categories include health system administrators and policymakers (n=30 across 5 countries), healthcare providers including physicians, nurses and community health workers (n=50 across 5 countries), AI researchers and health informatics specialists (n=25 across 5 countries) and patient advocacy representatives and civil society organization members (n=20 across 5 countries).

Sampling strategy within each stakeholder category prioritizes diversity across multiple dimensions including gender, professional experience, organizational affiliation and geographical location. Recruitment approaches utilize professional networks, organizational partnerships, snowball sampling and direct outreach through professional associations and advocacy organizations.

Data collection employs secure video conferencing platforms with provisions for telephone interviews where internet connectivity is limited. Interview duration ranges from 60-90 minutes with provisions for follow-up clarification interviews as needed. All interviews are recorded with participant consent and transcribed verbatim with quality assurance procedures including accuracy verification and de-identification protocols.

Interview guide development follows established qualitative research principles with questions addressing stakeholder experiences with AI in healthcare, perceptions of algorithmic bias and discrimination, organizational practices for AI evaluation and governance and recommendations for policy and practice improvements. Interview guides undergo cultural adaptation for each country context with input from local research partners and pilot testing with representative stakeholders.

Thematic analysis employs the framework approach with both deductive themes derived from theoretical frameworks and inductive themes emerging from data analysis [17]. Analysis procedures include multiple rounds of coding by independent researchers, regular team meetings to discuss emerging themes and stakeholder validation workshops to verify interpretation accuracy.

Healthcare system case studies

In-depth case studies provide detailed examination of algorithmic bias manifestation within specific healthcare system contexts, enabling comprehensive analysis of complex interactions between technological, organizational, social and policy factors. Case study methodology follows established principles for explanatory case study research with attention to construct validity, internal validity, external validity and reliability [18].

Case selection criteria prioritize countries with documented AI implementation in healthcare, availability of relevant stakeholders and data sources, established research partnerships and representation of diverse healthcare system characteristics. Selected countries include Nigeria (federal system with mixed public-private delivery), India (large-scale digital health initiatives), Kenya (mobile health innovation), Brazil (universal health system with AI integration) and Bangladesh (rural health program digitalization). Each case study employs multiple data sources to enable triangulation and comprehensive analysis. Administrative health data sources include electronic health records, health management information systems and AI system performance reports where available with appropriate privacy protections and data use agreements. Policy documentation encompasses national strategies, implementation guidelines, evaluation reports and regulatory frameworks.

Stakeholder interviews within each case study context involve approximately 25-30 participants representing diverse perspectives including government officials, healthcare managers, technology implementers, healthcare providers and community representatives. Semi-structured interview protocols address system-specific questions while maintaining consistency with the broader multi-stakeholder interview framework.

Technical system documentation includes AI system specifications, training data characteristics, validation procedures and deployment protocols obtained through partnerships with technology developers, healthcare organizations and government agencies. This documentation undergoes systematic analysis to identify potential bias sources, mitigation strategies and monitoring mechanisms.

Analytical procedures employ cross-case pattern matching and explanation building to identify common mechanisms and context-specific factors affecting algorithmic bias manifestation and impact [19]. Within-case analysis addresses temporal sequences, causal relationships and outcome patterns, while cross-case analysis identifies generalizable patterns and contextual variations.

Phase 3: Framework development and validation

Integrated analytical framework

The framework development phase synthesizes evidence from systematic review, policy analysis, stakeholder interviews and case studies to construct a comprehensive analytical framework for understanding and investigating algorithmic bias in LMIC healthcare contexts. Framework development employs established methodological approaches including framework synthesis and meta-ethnography to integrate diverse evidence sources while maintaining theoretical coherence and practical utility [20].

The synthesis process begins with systematic extraction of key findings, themes and concepts from each data source using standardized templates designed to capture both empirical findings and methodological insights. Extracted data undergoes thematic analysis to identify recurring patterns, contradictory findings and evidence gaps requiring further investigation.

Framework construction involves iterative development of conceptual models that integrate empirical findings with existing theoretical frameworks related to social determinants of health, health equity and algorithmic fairness. The framework development process includes regular consultation with subject matter experts, stakeholder validation workshops and peer review by international research collaborators.

The resulting analytical framework encompasses multiple components including theoretical models describing algorithmic bias pathways, operational definitions and measurement indicators, methodological guidelines for bias assessment and implementation strategies for different healthcare system contexts. Framework components undergo validation through application to case study data and stakeholder feedback processes.

Protocol validation and refinement

The final phase involves validation and refinement of research protocol components through pilot implementation in two additional LMIC contexts not included in the primary data collection phase. Pilot countries (Ghana and Vietnam) are selected to represent different geographical regions and healthcare system characteristics while providing opportunities for protocol testing in diverse contexts.

Pilot implementation focuses on protocol feasibility, cultural appropriateness and methodological robustness. Feasibility assessment addresses resource requirements, timeline adequacy, stakeholder accessibility and data availability within realistic implementation constraints. Cultural appropriateness evaluation involves systematic review of all protocol components with local stakeholders to identify potential cultural sensitivity issues, linguistic barriers and adaptation requirements.

Methodological validation employs established criteria including reliability assessment through inter-rater agreement analysis, validity evaluation through stakeholder feedback and expert review and practicality assessment through resource utilization and implementation timeline analysis. Pilot implementation generates systematic documentation of implementation challenges, adaptation requirements and methodological modifications.

Refinement processes incorporate stakeholder feedback, pilot implementation findings and expert consultation to optimize protocol components for broader implementation. Final protocol documentation includes comprehensive implementation guidelines, adaptation frameworks for different contexts, quality assurance procedures and dissemination strategies.

Implementation Considerations

Ethical framework

The ethical framework guiding this research protocol prioritizes justice, beneficence, respect for persons and cultural sensitivity while addressing specific ethical challenges associated with algorithmic bias research in LMIC contexts [21]. Justice considerations encompass fair participant selection, equitable benefit distribution and protection of vulnerable populations including rural communities, ethnic minorities and economically disadvantaged groups who may be disproportionately affected by algorithmic bias.

Beneficence principles require careful assessment of research risks and benefits with particular attention to potential harms from participating in algorithmic bias research, including privacy risks, stigmatization and unintended consequences of bias identification. Benefit maximization strategies include capacity building components, knowledge sharing commitments and collaborative research relationships that provide value to participating communities and institutions.

Respect for persons encompasses comprehensive informed consent procedures adapted for diverse cultural contexts and literacy levels, voluntary participation guarantees and participant autonomy protection throughout the research process. Consent procedures include clear explanation of research purposes, data use intentions, privacy protections and participant rights in culturally appropriate languages and formats.

https://doi.org/10.70844/jmhrp.2025.2.2.38

Cultural sensitivity provisions address diverse religious, cultural and social norms across research contexts through collaborative relationships with local communities, indigenous consultation processes where appropriate and cultural competency training for research staff. Community engagement strategies ensure meaningful participation in research design, implementation and dissemination processes.

Institutional Review Board (IRB) approval strategies involve comprehensive multi-country ethics review processes including institutional approvals in each participating country, international research ethics committee review and ongoing ethics monitoring throughout the research implementation. IRB applications include detailed protocols for vulnerable population protection, community benefit sharing and cultural sensitivity measures.

Data protection and privacy measures comply with international standards including relevant provisions of the European Union General Data Protection Regulation, national data protection laws in participating countries and international research data governance guidelines. Specific provisions address cross-border data transfer, long-term data storage and participant de-identification procedures.

Capacity building and partnership strategy

The capacity building strategy recognizes that sustainable research implementation and knowledge utilization require substantial investment in local research capacity, institutional partnerships and knowledge transfer mechanisms [22]. Partnership development prioritizes long-term collaborative relationships with academic institutions, government agencies, civil society organizations and healthcare systems in each participating country.

Local partnership requirements include formal agreements with established academic institutions in each study country, ensuring access to local expertise, institutional infrastructure and regulatory navigation support. Partner institutions must demonstrate relevant research experience, institutional capacity for ethical review and commitment to collaborative research relationships extending beyond the immediate project timeline.

Capacity building components encompass multiple dimensions including technical skills development, methodological training and institutional strengthening. Technical skills development addresses algorithmic bias assessment methods, qualitative research techniques, policy analysis approaches and research ethics procedures through comprehensive training programs, mentorship relationships and collaborative learning opportunities.

Research methodology training includes intensive workshops, online learning modules and practical implementation experiences designed to build local expertise in algorithmic bias research. Training curricula address study design, data collection methods, analytical techniques and dissemination strategies with specific attention to LMIC context adaptations and resource optimization approaches.

Institutional strengthening initiatives support research infrastructure development, ethical review capacity enhancement and policy engagement capabilities within partner institutions. These initiatives include equipment provision, software licensing, administrative support and collaborative planning for sustained research capacity beyond the immediate project period.

Knowledge transfer mechanisms ensure effective dissemination of research findings, methodological innovations and policy recommendations to diverse audiences including academic communities, policy makers, healthcare practitioners and civil society organizations. Dissemination strategies include peer-reviewed publications, policy briefs, stakeholder workshops and digital knowledge platforms.

Sustainability planning addresses long-term research collaboration development, funding diversification and institutional capacity maintenance through strategic planning processes, resource development initiatives and network strengthening activities. Partnership agreements include provisions for continued collaboration, shared intellectual property arrangements and mutual capacity building commitments.

Resource requirements and funding strategy

Comprehensive resource estimation for this 36-month research protocol totals approximately $2.5 million, including personnel costs, travel and accommodation, data collection expenses, capacity building activities and administrative overhead. Personnel costs encompass research staff salaries, consultant fees and stakeholder compensation across multiple countries with appropriate adjustments for local cost variations and currency fluctuations.

International travel requirements include research team coordination meetings, stakeholder interviews, case study site visits and dissemination activities with estimated costs incorporating pre-pandemic travel patterns and post-pandemic adaptation strategies including increased use of virtual meeting technologies where appropriate.

Technology and equipment costs address software licensing for data analysis, secure communication platforms, recording equipment and IT infrastructure support for virtual data collection and collaboration activities. These costs include provisions for backup systems, technical support and software updates throughout the project period.

Capacity building expenses encompass training workshop costs, educational material development, equipment provision for partner institutions and mentorship program implementation. These investments are designed to generate lasting benefits extending beyond the immediate research project through sustained local capacity enhancement.

Funding source diversification recognizes the importance of financial sustainability and stakeholder buy-in through multiple funding mechanisms. Primary funding targets include international development agencies such as the United States Agency for International Development, Department for International Development and Canadian International Development Agency with specific interests in digital health and health equity research.

Research council funding opportunities include competitive grants from bodies such as the well come Trust, Bill and Melinda Gates Foundation and Ford Foundation that prioritize global health research, digital inclusion and health equity initiatives. Academic funding sources encompass university research grants, international collaboration funds and multi-institutional partnership programs.

Cost-effectiveness optimization strategies maximize research impact per dollar invested through collaborative approaches, resource

https://doi.org/10.70844/jmhrp.2025.2.2.38

Journal of Medical Health Research and Psychiatry

sharing agreements and efficiency-focused implementation strategies. These approaches include joint training programs, shared data collection activities and coordinated dissemination efforts that reduce per-country implementation costs while maintaining methodological rigor.

Resource optimization measures address practical constraints in LMIC research environments through flexible implementation approaches, local procurement strategies and adaptive resource allocation mechanisms that can respond to changing conditions and unexpected challenges during research implementation.

Expected Outcomes and Impact

Methodological contributions

This research protocol will generate several significant methodological contributions to the emerging field of algorithmic bias research in global health contexts. The primary methodological output comprises a comprehensive, validated research protocol specifically designed for investigating algorithmic bias as a social determinant of health in LMIC healthcare systems. This protocol will provide the first systematic methodological framework addressing the unique challenges, opportunities and constraints characterizing LMIC research environments while maintaining scientific rigor and international comparability.

Standardized measurement tools and indicators represent a critical methodological contribution addressing current limitations in algorithmic bias assessment approaches. These tools will encompass both quantitative metrics for measuring bias manifestation across different algorithms and healthcare contexts and qualitative assessment instruments for capturing stakeholder perspectives, community experiences and contextual factors affecting bias impact. The measurement framework will include validation data demonstrating reliability and validity across diverse cultural and linguistic contexts.

Best practices for cross-cultural research implementation will provide practical guidance for researchers seeking to conduct algorithmic bias studies in diverse global contexts. These guidelines will address cultural adaptation procedures, linguistic translation protocols, community engagement strategies and ethical considerations specific to algorithmic bias research involving vulnerable populations. Implementation guidance will include cost-effective approaches for resource-constrained environments and quality assurance procedures for maintaining methodological consistency across different contexts.

Framework development for policy-relevant research design represents an important methodological innovation addressing the gap between academic research and policy application in algorithmic bias studies. This framework will provide guidance for designing research studies that generate evidence directly applicable to policy decisions while maintaining academic rigor and methodological soundness. The framework will include stakeholder engagement protocols, evidence synthesis approaches and dissemination strategies that maximize policy uptake and implementation.

Empirical contributions

The empirical evidence generated through this research protocol will address critical knowledge gaps regarding algorithmic bias prevalence, patterns and impacts in LMIC healthcare contexts. The systematic evidence mapping will provide the first comprehensive assessment of existing knowledge regarding algorithmic bias in LMIC healthcare systems, identifying evidence gaps, methodological limitations and research priorities for future investigation. This evidence base will inform policy decisions, funding priorities and research agenda development across multiple stakeholder communities.

Multi-stakeholder interview data will generate comprehensive understanding of how different actors within LMIC healthcare systems perceive, experience and respond to algorithmic bias concerns. This evidence will illuminate the complex social, cultural and organizational factors that influence bias manifestation and impact while providing insights into stakeholder priorities, concerns and recommendations for policy and practice improvements. Interview findings will inform the development of culturally appropriate interventions and stakeholder-responsive policy frameworks.

Healthcare system case studies will provide detailed empirical evidence regarding specific mechanisms through which algorithmic bias affects health outcomes in different LMIC contexts. Case study findings will document causal pathways, moderating factors and outcome patterns while identifying context-specific factors that influence bias manifestation and impact. This evidence will support the development of targeted interventions and context-appropriate policy recommendations.

Comparative analysis across multiple LMIC contexts will identify common patterns and contextual variations in algorithmic bias manifestation, enabling the development of generalizable findings while recognizing important contextual factors. Comparative findings will inform international cooperation strategies, best practice identification and policy harmonization efforts across different healthcare systems and regulatory environments.

The identification of vulnerable populations and high-risk contexts will provide critical evidence for targeting bias mitigation efforts and protecting communities most at risk of algorithmic discrimination. This evidence will support the development of vulnerability assessment frameworks, risk stratification approaches and targeted intervention strategies that prioritize equity and justice considerations.

Case studies demonstrating measurable bias impact on health outcomes will provide concrete evidence of the health equity implications of algorithmic bias in real-world healthcare delivery contexts. These demonstrations will support policy advocacy efforts, funding justification and stakeholder engagement initiatives while providing empirical validation of theoretical frameworks linking algorithmic bias to health equity outcomes.

Policy and practice implications

The research protocol will generate evidence-based recommendations for AI governance frameworks specifically tailored to LMIC healthcare contexts. These recommendations will address regulatory approaches, oversight mechanisms and governance structures that balance innovation promotion with bias mitigation and equity protection. Policy recommendations will include specific guidance for different healthcare system types, regulatory capacity levels and resource availability contexts.

Implementation guidelines for bias mitigation strategies will provide practical guidance for healthcare organizations, technology developers and policy makers seeking to address algorithmic bias in their specific contexts. These guidelines will encompass technical approaches for bias detection and mitigation, organizational policies and procedures and system-level interventions that address structural factors contributing to bias manifestation.

Capacity building frameworks will support healthcare systems in developing institutional capabilities for AI governance, bias assessment and equity monitoring. These frameworks will include training curricula, institutional development strategies and resource allocation guidance that enable sustainable capacity development within resource-constrained environments.

International cooperation strategies will facilitate knowledge sharing, best practice dissemination and collaborative approach development across different countries and healthcare systems. These strategies will include recommendations for multilateral initiatives, research collaboration frameworks and policy harmonization approaches that promote global health equity in the digital health era.

Limitations and Future Directions

Acknowledged limitations

This research protocol acknowledges several important limitations that may affect implementation feasibility and findings generalizability. Resource constraints represent a fundamental limitation, as comprehensive multi-country research requires substantial financial, human and technical resources that may not be readily available in all desired contexts. Limited funding availability may necessitate reduced country participation, shorter implementation timelines, or modified methodological approaches that could affect the comprehensiveness and comparability of findings.

Data availability variations across different LMIC contexts present significant methodological challenges that may limit analytical consistency and comparative analysis capabilities. Healthcare systems with limited electronic health records, inadequate data management systems, or restricted data access policies may require alternative data collection approaches or may be excluded from certain analytical components. These variations may affect the representativeness of findings and limit generalizability across different healthcare system contexts.

Cultural and linguistic barriers may affect data collection quality, stakeholder engagement effectiveness and findings interpretation accuracy despite comprehensive cultural adaptation procedures. Translation challenges, cultural sensitivity issues and communication difficulties may introduce systematic biases or reduce data quality in ways that are difficult to detect and control through standard quality assurance procedures.

The rapidly evolving technology landscape presents temporal limitations, as AI systems, regulatory frameworks and implementation approaches may change significantly during the research implementation period. Findings regarding specific technologies or policy approaches may become outdated before dissemination and emerging technologies may introduce new forms of bias that are not adequately addressed by existing methodological frameworks.

Regulatory and ethical approval challenges may affect implementation timelines, data collection approaches, or country participation in ways that compromise methodological consistency or analytical comprehensiveness. Variations in ethical review requirements, data protection regulations and research approval processes across different countries may necessitate methodological adaptations that affect comparability and generalizability.

Future research priorities

This research protocol establishes the foundation for an extensive future research agenda addressing algorithmic bias and health equity in global health contexts. Longitudinal studies examining bias impact over time represent a critical priority, as current understanding of algorithmic bias effects relies primarily on cross-sectional analyses that cannot capture temporal patterns, cumulative impacts, or long-term consequences of bias exposure.

Intervention effectiveness research represents another high-priority area, as evidence regarding successful bias mitigation strategies remains limited and context-specific. Future studies should employ rigorous experimental and quasi-experimental designs to evaluate the effectiveness of different bias mitigation approaches across diverse healthcare settings, population groups and technology applications.

Economic evaluation of bias mitigation strategies requires systematic investigation to inform resource allocation decisions and policy prioritization efforts. Cost-effectiveness analyses, budget impact assessments and return-on-investment calculations will support evidence-based decision-making regarding bias mitigation investments and intervention selection.

Scale-up and implementation science studies will address the challenge of translating successful bias mitigation interventions from pilot implementations to large-scale healthcare system adoption. These studies should examine implementation barriers, facilitators and adaptation requirements while developing frameworks for sustainable intervention maintenance and continuous improvement.

Population-specific research addressing bias impacts among particularly vulnerable groups, including indigenous populations, ethnic minorities, rural communities and individuals with disabilities, requires dedicated investigation to understand unique risk factors, impact patterns and intervention needs. These studies should employ participatory research approaches that center community perspectives and priorities.

Technology-specific investigations examining bias manifestation across different AI applications, including diagnostic algorithms, treatment recommendation systems and resource allocation tools, will provide targeted evidence for technology developers and healthcare implementers. These studies should address both technical and social factors affecting bias manifestation while developing application-specific mitigation strategies.

International comparative research examining bias patterns across different regions, healthcare systems and regulatory environments will support global policy development and international cooperation initiatives. These studies should identify transferable best practices while recognizing important contextual factors that affect intervention effectiveness and policy appropriateness.

https://doi.org/10.70844/jmhrp.2025.2.2.38

Journal of Medical Health Research and Psychiatry

Conclusion

This research protocol presents the first comprehensive methodological framework specifically designed for investigating algorithmic bias as a social determinant of health in low- and middle-income country healthcare contexts. By addressing critical evidence gaps and methodological limitations in current research approaches, this protocol establishes the foundation for systematic, rigorous and culturally appropriate investigation of algorithmic bias phenomena across diverse global health settings.

The multi-phase, mixed-methods approach provides balanced consideration of methodological rigor and practical implementation feasibility while maintaining sensitivity to the complex cultural, economic and healthcare system contexts that characterize LMIC environments. Through systematic evidence mapping, comprehensive stakeholder engagement, detailed case study analysis and collaborative framework development, this protocol generates the methodological tools and empirical evidence necessary for evidence-based policy development and effective bias mitigation strategies.

The significance of this work extends beyond immediate research contributions to encompass broader implications for global health equity, international development and technology governance in healthcare. As AI technologies become increasingly integrated into healthcare delivery systems worldwide, ensuring equitable access to beneficial technologies while protecting vulnerable populations from algorithmic discrimination represents a fundamental challenge for global health and social justice.

The collaborative approach emphasized throughout this protocol reflects recognition that addressing algorithmic bias requires sustained international cooperation, knowledge sharing and capacity building initiatives that transcend traditional boundaries between research, policy and practice communities. By establishing partnerships across multiple countries, stakeholder groups and disciplinary perspectives, this research creates the foundation for long-term collaborative efforts addressing algorithmic bias and health equity.

Implementation of this research protocol requires commitment from multiple stakeholder communities including academic institutions, funding organizations, policy makers, healthcare leaders and civil society organizations. The 36-month timeline provides sufficient duration for comprehensive investigation while maintaining momentum and stakeholder engagement throughout the research process.

The expected outcomes encompass both immediate methodological contributions and longer-term empirical evidence that will inform policy decisions, guide intervention development and support capacity building initiatives across diverse LMIC contexts. By providing standardized research approaches, validated measurement tools and evidence-based policy recommendations, this protocol contributes to the development of more equitable, effective and ethical AI implementation in global health contexts.

Future research priorities emerging from this protocol encompass longitudinal impact studies, intervention effectiveness research, economic evaluations and implementation science investigations that will further advance understanding and practice in this rapidly evolving field. The protocol establishes the methodological foundation for sustained research programs addressing algorithmic bias and health equity across diverse global contexts.

This protocol represents a call for international collaboration and coordinated action to ensure that the digital health transformation occurring across low- and middle-income countries promotes rather than undermines health equity and social justice. Through systematic investigation, evidence-based policy development and collaborative capacity building, the global health community can work together to harness the benefits of AI technologies while protecting vulnerable populations from algorithmic discrimination.

The commitment to open science principles and knowledge sharing embedded throughout this protocol ensures that research findings, methodological innovations and policy recommendations will be freely available to researchers, policy makers and practitioners worldwide. This approach maximizes the potential impact of research investments while promoting transparency, accountability and collaborative learning across the global health community.

Competing Interests

The authors report no conflicts of interest in this work.

References

1. Schwalbe N, Wahl B (2020) Artificial intelligence and the future of global health. The Lancet 395(10236): 1579-1586. [GoogleScholar]
2. Alami H, Gagnon MP, Fortin JP (2020) Digital health and the challenge of health systems transformation. mHealth 3:31. [Crossref] [GoogleScholar]
3. Wahl B, Cossy Gantner A, Germann S, Schwalbe NR (2018) Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings? BMJ Global Health 3(4): e000798.
4. Gwagwa A, Kraemer-Mbula E, Rizk N, Rutenberg I, De Beer J (2021) Artificial intelligence (AI) governance in Africa: A participatory approach. Proceedings of the 2021 ACM Conference on Fairness, Accountability and Transparency, 329-342.
5. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH (2018) Ensuring fairness in machine learning to advance health equity. Annals of Internal Medicine 169(12): 866-872. [Crossref] [GoogleScholar]
6. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, et al. (2021) Ethical machine learning in healthcare. Annual Review of Biomedical Data Science 4: 123-144. [Crossref] [GoogleScholar]
7. Reddy S, Allan S, Coghlan S, Cooper P (2020) A governance model for the application of AI in health care. J Am Med Inform Assoc 27(3), 491-497. [Crossref] [GoogleScholar]
8. Barocas S, Hardt M, Narayanan A (2019) Fairness and machine learning: Limitations and opportunities. MIT Press. [GoogleScholar]
9. Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. Science 366(6464): 447-453. [Crossref] [GoogleScholar]
10. Veinot TC, Mitchell H, Ancker JS (2018) Good intentions are not enough: How informatics interventions can worsen inequality. JAMIA 25(8): 1080-1088. [Crossref] [GoogleScholar]
11. Marmot M, Wilkinson RG (2005) Social determinants of health. Oxford University Press.
12. Benjamin R (2019) Race after technology: Abolitionist tools for the new Jim Code. Polity Press.
13. Zou J, Schiebinger L (2018) AI can be sexist and racist—It's time to make it fair. Nature 559(7714): 324-326. [Crossref] [GoogleScholar]

https://doi.org/10.70844/jmhrp.2025.2.2.38

Journal of Medical Health Research and Psychiatry

14. Larson DB, Magnus DC, Lungren MP, Shah NH, Langlotz CP (2020) Ethics of using and sharing clinical imaging data for artificial intelligence: A proposed framework. Radiology 295(3): 675-682. [Crossref] [GoogleScholar]
15. Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, et al. (2018) High-quality health systems in the Sustainable Development Goals era: Time for a revolution. The Lancet Global Health 6(11): e1196-e1252. [Crossref] [GoogleScholar]
16. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H,at et al. (2018) PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Annals of Internal Medicine 169(7): 467-473. [Crossref] [GoogleScholar]
17. Gale NK, Heath G, Cameron E, Rashid S, Redwood S (2013) Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology 13: 117. [Crossref] [GoogleScholar]
18. Yin RK (2018) Case study research and applications: Design and methods (6th ed.). Sage Publications.
19. Miles MB, Huberman AM, Saldaña J (2014) Qualitative data analysis: A methods sourcebook (3rd ed.). Sage Publications.
20. Carroll C, Booth A, Cooper K (2013) A worked example of "best fit" framework synthesis: A systematic review of views concerning the taking of some potential chemopreventive agents. BMC Med Res Methodol 13: 117. [Crossref]
21. Emanuel EJ, Wendler D, Grady C (2000) What makes clinical research ethical? JAMA 283(20): 2701-2711. [Crossref] [GoogleScholar]
22. Whitworth J, Sewankambo NK, Snewin VA (2010) Improving implementation: Building research capacity in maternal, neonatal and child health in Africa. PLoS Medicine 7(7): e1000299. [Crossref] [GoogleScholar]