Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Governance Determinants of AI-Enabled Mental Health Systems in LMICs

Effiong Ndarake Effiong

effiongneffiong@gmail.com

Nasarawa State University

Daniel Mairafi Gimbason

Nasarawa State University

Yakubu Boyi Ngwai

Nasarawa State University

Article

Keywords:

Posted Date: March 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8981586/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Digital and AI-enabled tools are increasingly integrated into mental health systems in low-and middle-income countries to address treatment gaps. While clinical effectiveness and ethical principles have received attention, structural governance determinants influencing equity, accountability, and system integration remain less systematically examined. We conducted a structured policy and literature analysis (2015–2025) of AI-enabled mental health deployments in LMICs, analyzing 78 peer-reviewed publications and policy documents using best-fit framework synthesis. Governance-relevant patterns were coded across accountability mechanisms, data governance, infrastructural inclusion, and cultural validation. Across cases, four recurring patterns emerged: limited post-deployment monitoring of algorithmic performance; data governance framed as privacy compliance rather than structural stewardship; infrastructural access constraints affecting participation; and variable attention to cultural validation beyond translation. These findings suggest implementation outcomes are shaped not only by clinical effectiveness but by governance arrangements. Clarifying these governance domains may support more systematic integration of digital tools within mental health systems, though empirical validation in prospective studies is warranted.

Introduction

Mental health conditions contribute substantially to the global burden of disease, with persistent treatment gaps across many low- and middle-income countries (LMICs) [1, 2]. Digital and AI-enabled mental health interventions have been proposed as scalable approaches to expanding access and addressing workforce shortages [3–6]. Evidence suggests that digital psychological interventions can produce clinically meaningful symptom reductions in selected LMIC contexts [4, 6, 7].

However, implementation outcomes remain heterogeneous. Increasingly, digital health research recognizes that infrastructural capacity, regulatory environments, health system organization, and sociocultural context shape both effectiveness and reach [8–11]. Within global mental health, there is growing emphasis on systems strengthening, equity-sensitive implementation, and structural determinants of care delivery [12, 13].

Although normative ethical principles for artificial intelligence in health have been articulated at global level [14], less attention has been directed toward how governance arrangements operate in practice within LMIC mental health systems. Governance, in this context, extends beyond high-level ethical commitments and refers to the institutional mechanisms that structure accountability, data stewardship, infrastructural inclusion, and cultural validation across the lifecycle of AI-enabled systems. Prior analyses of AI ethics in healthcare have noted that translation from principle to operational governance mechanisms remains uneven across contexts [3]. As AI-enabled mental health tools move from pilot implementation to population-scale deployment, variability in governance arrangements may materially influence equity, accountability, and system resilience.

This study aims to identify recurring governance determinants across documented AI-enabled mental health deployments in LMICs and to organize them within a systems-oriented analytical framework. Rather than proposing regulatory prescriptions, the analysis seeks to clarify how governance functions in practice and to provide a structured vocabulary for examining how AI systems integrate within broader mental health ecosystems.

Methods

Study Design

A structured policy and literature analysis was conducted using best-fit framework synthesis [15], consistent with health policy analysis methodologies [16, 17].

Search Strategy

Peer-reviewed articles and official policy documents (January 2015 - December 2025) were identified through PubMed, Web of Science, and WHO IRIS. Inclusion criteria:

- AI-enabled or algorithm-based mental health intervention
- LMIC setting
- Discussion of implementation, governance, equity, or regulation

Seventy-eight documents met inclusion criteria.

Analysis

Two authors independently reviewed and coded included documents. Initial coding categories were informed by prior governance literature; accountability, data governance, infrastructure, cultural validity. Additional subcategories were developed inductively. Coding discrepancies were resolved through discussion. Cross-case comparison assessed whether identified patterns recurred across distinct LMIC contexts.

This review was conducted using a structured search and analytic protocol consistent with PRISMA 2020 reporting principles for evidence synthesis. Search terms combined concepts related to artificial intelligence, algorithmic or digital mental health interventions, governance, implementation, equity, and LMIC settings. Searches were performed across PubMed, Web of Science, and WHO IRIS databases for publications dated January 2015 through December 2025. After removal of duplicates, titles and abstracts were screened independently by two authors against predefined inclusion criteria. Full texts were assessed for eligibility where relevance was uncertain. Disagreements were resolved through consensus discussion. A flow diagram documenting identification, screening, eligibility, and inclusion processes is provided in the Supplementary Materials. Data extraction was conducted using a structured

template aligned with governance domains derived from best-fit framework synthesis methodology [15], allowing both deductive coding based on established governance constructs and inductive refinement based on cross-case comparison.

Results

Algorithmic Accountability

Several studies described evaluation during pilot or research phases [6, 18], but fewer detailed structured post-deployment monitoring of algorithmic performance across demographic subgroups. Engagement disparities by gender and socioeconomic status were reported in some digital mental health programs [19, 20], yet formalized auditing mechanisms were not consistently described. This pattern reflects broader observations that while clinical outcomes receive attention, systems for continuous accountability remain underdeveloped [3].

Data Stewardship

Most reports referenced compliance with ethical review and data protection regulations [14, 20]. However, documentation of mechanisms for ongoing oversight of data reuse, retraining, or cross-border data flows was limited. Global health governance scholarship has noted similar patterns in other digital health domains [21]. The emerging emphasis on digital sovereignty including data residency, jurisdictional control, and community oversight has not yet translated into routine governance requirements for AI mental health systems in LMICs [22].

Infrastructure Inclusion

Digital divide literature highlights persistent disparities in device ownership and connectivity in LMICs [8, 23]. Several mental health implementation reports acknowledged connectivity barriers affecting participation [9, 10]. Infrastructure constraints were generally discussed descriptively rather than embedded within governance requirements, despite evidence that network-based AI systems risk reproducing existing inequities when infrastructural access is not actively governed [24].

Cultural Validation

Cross-cultural psychiatry research demonstrates variability in symptom presentation and idioms of distress [25, 26]. While translation and contextual adaptation were frequently described, fewer studies reported formal evaluation of algorithmic validity across cultural subgroups. Participatory design was documented in some cases [5], though institutionalization varied. Recent work on culturally grounded frameworks, such as Ubuntu-guided AI for mental health in African contexts, highlights pathways for deeper epistemic inclusion [27]. However, such approaches remain exceptional rather than institutionalized.

Analytic Organization of Governance Domains

Across cases, governance determinants clustered into four domains, see Table 1

Table 1
Governance domains identified across AI-enabled mental health implementations in LMICs

Governance domain	Empirical pattern identified across cases	Representative sources
Algorithmic accountability	Limited post-deployment monitoring; engagement disparities reported but not routinely audited across subgroups	[3, 6, 18, 19]
Data stewardship	Privacy compliance emphasized; limited documentation of structural data control, retraining governance, or cross-border oversight	[14, 20–22]
Infrastructure inclusion	Access constraints acknowledged descriptively but rarely embedded as governance requirements	[8–10, 23, 24]
Cultural validation	Translation and contextual adaptation common; formal cross-cultural algorithmic validation less frequently reported	[5, 25–27]

Caption

Governance domains derived through cross-case synthesis of AI-enabled mental health implementations in LMICs (2015–2025). Domains are analytical constructs intended to clarify recurring governance dimensions rather than to prescribe regulatory standards.

These domains are derived from cross-case patterns and intended to clarify governance dimensions within global mental health systems research rather than to prescribe regulatory standards.

Discussion

This analysis suggests that AI-enabled mental health systems in LMICs operate within governance environments that vary in structure, maturity, and institutional clarity. While global ethical guidance for AI in health exists [14], implementation documentation indicates uneven operationalization across four interrelated domains: algorithmic accountability, data stewardship, infrastructure inclusion, and cultural validation.

Across cases, monitoring of algorithmic performance frequently occurred during pilot or trial phases but was less consistently described in post-deployment contexts. Engagement disparities across gender and socioeconomic strata were reported in some settings [19, 20], yet formalized mechanisms for continuous auditing or subgroup performance review were not routinely documented. Diffuse responsibility without clearly allocated institutional oversight risks weakening accountability across the AI system lifecycle [3].

Data governance was commonly framed in terms of compliance with ethical review procedures and data protection regulations [14, 20]. However, mechanisms addressing longer-term stewardship including retraining governance, secondary data use, and cross-border data flows were less consistently articulated. Similar structural gaps have been observed in broader digital health governance analyses [21]. As digital sovereignty debates gain prominence in LMIC policy discourse [22], the absence of explicit governance mechanisms within AI mental health deployments may have implications for sustainability and public trust.

Infrastructure inclusion emerged as a recurrent constraint. Device ownership, connectivity disparities, and digital literacy continue to shape participation in LMIC contexts [8, 23]. While several implementation reports acknowledged these barriers [9, 10], infrastructural access was rarely embedded as a formal governance requirement. Network-based AI systems, particularly those dependent on continuous data exchange, may risk reproducing existing inequities if infrastructural determinants are not explicitly governed [24].

Cultural validation represented a fourth domain of variability. Cross-cultural psychiatry demonstrates heterogeneity in idioms of distress and symptom expression [25, 26]. Translation and contextual adaptation were commonly described; however, systematic evaluation of algorithmic validity across cultural subgroups was less frequently reported. Participatory and human centred design approaches were documented in selected initiatives [5], and emerging frameworks propose more epistemically grounded models for AI development in African contexts [27]. Nevertheless, such approaches remain unevenly institutionalized.

Taken together, these findings indicate that governance arrangements may shape not only regulatory compliance but also distributive equity, performance stability, and legitimacy of AI-enabled mental health systems. The four domains identified here provide a structured analytic vocabulary for examining how responsibility is distributed across the AI lifecycle from model development and deployment to monitoring, updating, and integration within existing health systems. This framing aligns with calls in global mental health for equity-sensitive implementation science [12, 13] and with scholarship emphasizing structural determinants of knowledge production and governance in global health [28].

Positioning governance determinants within a systems-oriented framework enables more explicit examination of how technical architectures, institutional arrangements, and sociocultural contexts interact. Future empirical research could test associations between governance characteristics and measurable implementation outcomes, including service reach, sustained engagement, subgroup performance equity, and population-level impact. Such analyses would strengthen the evidence base for integrating AI-enabled tools into LMIC mental health systems in ways that are both technically robust and institutionally resilient.

Limitations

This analysis draws on published and officially documented sources and therefore may not capture informal, emergent, or undocumented governance practices operating within implementation contexts. As a structured policy and literature synthesis, it does not assess clinical effectiveness or directly measure implementation outcomes. The governance domains identified are analytically derived and require empirical validation in prospective implementation and comparative systems research. Finally, the analysis foregrounds institutional governance mechanisms and does not comprehensively examine political economy dynamics such as financing structures, power asymmetries, and incentive architectures that may shape governance feasibility and sustainability.

Conclusion

AI-enabled mental health systems are expanding within LMIC contexts. Evidence indicates that governance determinants related to accountability, data stewardship, infrastructure, and cultural validation influence how these systems function within broader health systems. Clarifying these domains contributes to global mental health systems research and provides a basis for future empirical investigation.

Declarations

Authors' Contributions

Effiong Ndarake Effiong: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing of Original Draft, Visualization, Project Administration. He led the study's design and execution.

Gimbason Mairafi Daniel: Methodology, Validation, Resources, Writing, Review & Editing and Supervision.

Yakubu B. Ngwai: Conceptualization, Methodology, Resources, Supervision, Writing, Review & Editing.

All authors meet the ICMJE authorship criteria, have read and approved the final manuscript, and agree to be accountable for all aspects of the work.

Ethics approval and consent to participate

Not Applicable

Consent for publication

Not applicable.

Availability of data and materials

Not Applicable

Competing interests: None declared. All authors have no competing interests to declare

Declaration of interests

The authors declare that they have no competing interests.

Funding

No specific funding was received for this study.

Acknowledgements

We thank the Management and Staff of Nasarawa State University, especially the Director of Global Health and Infectious Disease Institute (GHIDI), Prof. Adamu Ishiaku Akyala and all staff and students of Efficacy International School, Kaduna, Nigeria and its dedicated research directorate the Center for Applied Health Research and Development. The author used ChatGPT 5.0 for limited language refinement and editorial support; all conceptual framing, argumentation, and analysis are the author's own.

Author Declaration

We confirm that this manuscript represents original work, has not been published elsewhere, and is not currently under consideration by any other journal. All authors have read and approved the final manuscript and meet the ICMJE authorship criteria. The authors declare that they have no competing interests and received no specific funding for this study.

Thank you for considering our submission. We look forward to your positive decision regarding this timely and essential topic in global digital health governance.

References

1. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry* 9, 137–150 (2022).
2. Moitra, M. *et al.* The global gap in treatment coverage for major depressive disorder in 84 countries. *Lancet Psychiatry* 9, 137–150 (2022).
3. Morley, J., Machado, C. C. V., Burr, C., Cowls, J., Joshi, I., Taddeo, M. & Floridi, L. The ethics of AI in health care: A mapping review. *Soc. Sci. Med.* 260, 113172 (2020).
4. Naslund, J. A. *et al.* Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review. *Lancet Psychiatry* 4, 486–500 (2017).

5. Hall, J. *et al.* Sustainable Technology for Adolescents and Youth to Reduce Stress (STARS): a WHO transdiagnostic chatbot for distressed youth. *World Psychiatry* 21, 156–157 (2022).
6. Cuijpers, P. *et al.* Effects of a digital psychological intervention on Syrian refugees in Lebanon: a randomised controlled trial. *Lancet Psychiatry* 9, 564–574 (2022).
7. Firth, J. *et al.* The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 16, 287–298 (2017).
8. Veinot, T. C., Mitchell, H. & Ancker, J. S. Good intentions are not enough: how informatics interventions can worsen inequality. *J. Am. Med. Inform. Assoc.* 25, 1080–1088 (2018).
9. Weiss, D. *et al.* The digital divide in mental health: a systematic review. *Soc. Psychiatry Psychiatr. Epidemiol.* 57, 1531–1545 (2022).
10. Li, T., Wandella, A., Gomer, R. & Al-Mafazy, M. H. Operationalising health data governance for AI innovation in low-resource government health systems. *Data Policy* 6, e14 (2024).
11. Peters, D. H., Adam, T., Alonge, O., Agyepong, I. A. & Tran, N. Implementation research: what it is and how to do it. *BMJ* 347, f6753 (2013).
12. Damschroder, L. J., Reardon, C. M., Widerquist, M. A. O. & Lowery, J. The updated Consolidated Framework for Implementation Research (CFIR 2.0). *Implement. Sci.* 17, 75 (2022).
13. Abimbola, S. The uses of knowledge in global health. *BMJ Glob. Health* 6, e005802 (2021).
14. World Health Organization. *Ethics and Governance of Artificial Intelligence for Health*. (WHO, Geneva, 2021).
15. Carroll, C., Booth, A. & Cooper, K. A worked example of “best fit” framework synthesis. *BMC Med. Res. Methodol.* 11, 29 (2011).
16. Walt, G., Shiffman, J., Schneider, H., Murray, S. F., Brugha, R. & Gilson, L. ‘Doing’ health policy analysis: methodological and conceptual reflections. *Health Policy Plan.* 23, 308–317 (2008).
17. Gilson, L. (ed.) *Health Policy and Systems Research: A Methodology Reader*. (WHO, Geneva, 2012).
18. Torous, J. *et al.* The growing field of digital psychiatry. *World Psychiatry* 20, 318–335 (2021).
19. Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366, 447–453 (2019).
20. Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K. & Ghassemi, M. Ethical machine learning in health care. *Annu. Rev. Biomed. Data Sci.* 4, 123–144 (2021).
21. Taylor, L. & Broeders, D. In the name of development: power, profit and the datafication of the global South. *Geoforum* 64, 229–237 (2015).
22. African Union Commission. *African Union Data Policy Framework*. (African Union, Addis Ababa, 2022).
23. Robinson, L. *et al.* Digital inequalities 2.0: legacy inequalities in the information age. *First Monday* 25(7) (2020).
24. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ Digit Med. 2023;6(1):120. doi: 10.1038/s41746-023-00873-0.

25. Kohrt, B. A. *et al.* Cultural concepts of distress and psychiatric disorders: literature review and research recommendations. *Int. J. Epidemiol.* 43, 365–406 (2014).
26. Kirmayer, L. J. Cultural psychiatry in historical perspective. In: Bhugra, D. & Bhui, K. (eds.) *Textbook of Cultural Psychiatry*, 3–19 (Cambridge Univ. Press, 2007).
27. Abimbola S, Pai M. Will global health survive its decolonisation? Lancet. 2020;396(10263):1627–1628. doi: 10.1016/S0140-6736(20)32417-X.
28. Bhakuni, H. & Abimbola, S. Epistemic injustice in academic global health. *Lancet Glob. Health* 9, e1465–e1470 (2021).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterialsGovernanceDeterminants.pdf