COLOMBO SCIENTIFIC PUBLISHING

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM
E-ISSN: 2536-7919
P-ISSN: 2536-7900

PAGE NO: 52-56

Artificial Intelligence-Enabled Strategies for Enhancing Equity, Diversity, and Inclusion in Randomized Clinical Trials: Integrating Digital Health Infrastructure, Epidemiological Disparities, and Regulatory Innovation

Dr. Elena Marquez
Department of Health Policy and Biomedical Informatics University of Barcelona, Spain

ARTICLE INFO

Article history:

Submission: February 01, 2026
Accepted: February 15, 2026
Published: February 22, 2026
VOLUME: Vol.11 Issue 02 2026

Keywords:
artificial intelligence; randomized clinical trials; health equity; digital health infrastructure; federated learning; epidemiological disparities; regulatory innovation

ABSTRACT

Persistent inequities in randomized clinical trials (RCTs) compromise the generalizability, ethical integrity, and scientific validity of biomedical research. Disparities in enrollment by race, ethnicity, age, and socioeconomic status reflect broader structural inequities in healthcare, including those observed in mortality outcomes, chronic disease burden, oncology genomics, and infectious disease sequelae. The rapid digitization of healthcare and advances in artificial intelligence (AI) and machine learning (ML) offer unprecedented opportunities to redesign clinical trial ecosystems.

A structured, integrative research synthesis was conducted based on peer-reviewed studies, governmental guidance, and technological frameworks. The methodology synthesizes evidence on mortality disparities, disease epidemiology, genomic heterogeneity, electronic health record (EHR) adoption, cloud computing, federated learning, and regulatory digital health policies. These domains were analytically mapped onto stages of the clinical trial lifecycle.

The findings indicate that AI-driven eligibility optimization, predictive recruitment modeling, EHR-based phenotyping, federated multi-institutional collaboration, and cloud-enabled trial management can substantially mitigate structural barriers to participation. When aligned with regulatory digital innovation frameworks and ethical safeguards, AI systems can address underrepresentation in oncology, cardiometabolic disorders, infectious disease outcomes, and vaccine research. However, algorithmic bias, data fragmentation, and unequal digital access remain critical challenges.

AI/ML technologies, embedded within robust regulatory and ethical governance structures, can reconfigure RCTs toward structural inclusivity. Achieving equity requires not only technical innovation but also deliberate integration of epidemiological knowledge, cost-sensitive trial design, and digital infrastructure harmonization. Future research must empirically validate these frameworks in multi-national contexts.

INTRODUCTION

Randomized clinical trials (RCTs) remain the cornerstone of evidence-based medicine, providing the highest level of causal inference for therapeutic efficacy and safety. Yet despite their methodological rigor, RCTs have long been criticized for failing to adequately represent the populations most affected by disease. Structural inequities in enrollment across racial and ethnic groups, women, older adults, and socioeconomically disadvantaged populations undermine the external validity and ethical legitimacy of biomedical innovation (Flores et al., 2021). These disparities are not merely statistical artifacts; they are deeply intertwined with broader patterns of morbidity, mortality, and social stratification.

Evidence of persistent health inequities is well established. Within the Veterans Health Administration, racial and ethnic minority groups exhibit mortality disparities shaped by complex interactions among access, comorbidities, and social determinants (Peterson et al., 2018). In oncology, biologically distinct disease trajectories have been observed across populations, including racial and ethnic differences in clonal hematopoiesis and outcomes among patients with multiple

myeloma (Peres et al., 2022), as well as differential androgen receptor cistrome activity in prostate cancer among African American men (Berchuck et al., 2022). These findings demonstrate that underrepresentation in trials is not only an ethical issue but also a scientific liability, potentially obscuring biologically relevant variation.

Cardiometabolic disorders provide another lens into inequity. The epidemiology of obesity and diabetes reveals disproportionate burdens in marginalized communities, with downstream cardiovascular complications amplifying risk (Bhupathiraju & Hu, 2016). Structural determinants such as residential segregation further compound disparities in obesity and diabetes prevalence (Kershaw & Pender, 2016). Public health surveillance confirms persistent inequities in diabetes outcomes (Centers for Disease Control and Prevention, 2021). When RCTs evaluating novel therapeutics fail to reflect these epidemiological realities, clinical guidance risks being misaligned with population needs.

The COVID-19 pandemic intensified scrutiny of representativeness. Acute kidney injury in hospitalized COVID-19 patients demonstrated heterogeneous risk patterns influenced by comorbidities and systemic inequities (Hirsch et al., 2020). Simultaneously, evaluations of vaccine clinical trials revealed inconsistent inclusion of racial/ethnic minorities, women, and older individuals (Flores et al., 2021). These findings underscore the urgency of rethinking trial design in light of population-level vulnerability.

Beyond ethical and epidemiological considerations, economic pressures within pharmaceutical innovation shape trial structures. Estimates of research and development (R&D) costs for new pharmaceuticals have escalated dramatically (DiMasi et al., 2016). Clinical trials represent a major cost driver, influenced by complexity, enrollment rates, monitoring requirements, and failure risk (Sertkaya et al., 2016). High failure rates and inefficiencies further strain the system (Fogel, 2018), while broader transformations in drug development highlight the evolving face of innovation (Kaitin, 2010). Underrepresentation can exacerbate inefficiencies by necessitating post-marketing studies or failing to detect population-specific safety signals.

Digital transformation offers a potential paradigm shift. Electronic health record (EHR) adoption among office-based physicians has reached substantial penetration (HealthIT.gov, 2023), generating vast data reservoirs. Advances in deep learning in healthcare enable sophisticated pattern recognition and predictive analytics (Esteva et al., 2021). Unsupervised representation learning from EHRs has demonstrated capacity to predict future patient outcomes (Miotto et al., 2016). Cloud computing infrastructures facilitate scalable data processing (Rajan et al., 2021), while federated learning enables multi-institutional collaboration without direct data sharing (Sheller et al., 2020). Regulatory agencies have concurrently articulated digital innovation strategies (FDA, 2023; European Medicines Agency, 2021), recognizing the inevitability of computational integration into clinical research.

Within this context, AI/ML-based strategies for enhancing equity, diversity, and inclusion (EDI) in RCTs represent both a technological and moral frontier (Abbidi & Sinha, 2026). Yet the literature remains fragmented. While individual studies address disparities, digital tools, or economic pressures, a comprehensive integrative framework aligning epidemiological inequities with AI-enabled trial redesign remains underdeveloped.

This article addresses that gap. It synthesizes evidence across epidemiology, oncology genomics, cardiometabolic disease, infectious disease outcomes, digital health infrastructure, clinical trial economics, and regulatory guidance to develop a theoretical and operational model for AI-driven EDI enhancement in RCTs. By situating technological tools within the structural realities of health disparities and cost constraints, this study advances a holistic paradigm for equitable biomedical innovation.

METHODOLOGY

This research employs a structured integrative synthesis methodology designed to unify heterogeneous domains of scholarship into a coherent analytical framework. Rather than conducting primary empirical experimentation, the study systematically integrates findings from peer-reviewed biomedical research, public health epidemiology, health economics, digital health informatics, and regulatory guidance documents. The methodological approach proceeds through five interrelated phases: conceptual mapping, thematic clustering, lifecycle alignment, normative integration, and critical synthesis.

First, conceptual mapping involved extracting core constructs from the referenced literature. These constructs included mortality disparities (Peterson et al., 2018), genomic heterogeneity in oncology (Peres et al., 2022; Berchuck et al., 2022), cardiometabolic epidemiology (Bhupathiraju & Hu, 2016; Kershaw & Pender, 2016; CDC, 2021), infectious disease complications (Hirsch et al., 2020), trial cost drivers (DiMasi et al., 2016; Sertkaya et al., 2016), trial failure determinants (Fogel, 2018), and digital transformation tools including deep learning, EHR analytics, cloud computing,

and federated learning (Esteva et al., 2021; Miotto et al., 2016; Rajan et al., 2021; Sheller et al., 2020). Regulatory frameworks for digital innovation were also incorporated (FDA, 2023; European Medicines Agency, 2021).

Second, thematic clustering organized these constructs into four macro-domains: (1) Epidemiological Disparities and Biological Variation; (2) Economic and Structural Dynamics of Clinical Trials; (3) Digital Health Infrastructure and AI Capabilities; and (4) Regulatory and Governance Ecosystems. Each macro-domain was analyzed for its implications on trial representativeness.

Third, lifecycle alignment mapped these macro-domains onto discrete phases of the RCT lifecycle: trial design, site selection, participant recruitment, data monitoring, analysis, and post-market surveillance. This alignment enabled identification of intervention points where AI/ML systems could theoretically mitigate inequities.

Fourth, normative integration incorporated ethical and equity considerations into each lifecycle phase. Drawing from evidence of disparities and digital governance guidance, the framework assessed potential benefits and risks of AI implementation, including algorithmic bias and digital exclusion.

Finally, critical synthesis generated a unified theoretical model articulating how AI-enabled infrastructures, when embedded within equitable governance structures, can restructure RCT processes to enhance EDI.

The methodological strength of this approach lies in its interdisciplinarity and systemic perspective. By refusing reductionism-whether technological determinism or purely sociological critique-the analysis recognizes that equity in RCTs emerges from interactions among epidemiology, economics, digital systems, and regulation. Limitations include reliance on existing literature and absence of primary quantitative validation; however, the theoretical rigor and comprehensive integration provide a robust foundation for future empirical investigation.

RESULTS

The integrative synthesis yields a multi-dimensional framework for AI-enabled equity in RCTs. Results are presented descriptively across lifecycle stages.

AI-Driven Trial Design and Eligibility Optimization

Traditional eligibility criteria often exclude individuals with comorbidities, advanced age, or socioeconomic complexity, inadvertently marginalizing populations disproportionately affected by disease (Fogel, 2018). AI-driven modeling using EHR-derived phenotypes can simulate heterogeneous populations and predict safety outcomes across broader inclusion parameters (Miotto et al., 2016). By incorporating epidemiological data on obesity, diabetes, and cardiovascular risk (Bhupathiraju & Hu, 2016), algorithms can forecast adverse event probabilities in historically excluded groups, supporting evidence-based expansion of criteria.

In oncology, genomic heterogeneity identified among racial and ethnic groups (Peres et al., 2022; Berchuck et al., 2022) can inform stratified trial arms. AI systems analyzing multi-omic data can detect population-specific biomarkers, ensuring that trial endpoints capture biologically relevant diversity.

Predictive Recruitment and Site Selection

Underrepresentation often reflects geographic and structural barriers linked to residential segregation (Kershaw & Pender, 2016). AI-driven geospatial analytics integrating EHR adoption data (HealthIT.gov, 2023) and demographic mapping can identify underserved regions with high disease prevalence but low trial penetration. Cloud-based platforms facilitate real-time recruitment tracking (Rajan et al., 2021), enabling adaptive site allocation.

Predictive models trained on historical enrollment data can anticipate dropout risk and tailor engagement strategies. Such models must be audited for bias to prevent reinforcement of existing disparities.

Federated Multi-Institutional Collaboration

Data silos impede comprehensive representation. Federated learning allows institutions to collaboratively train models without sharing raw patient data (Sheller et al., 2020). This is particularly relevant for capturing rare disease manifestations in minority populations. Federated frameworks preserve privacy while enabling algorithmic generalization across diverse cohorts.

Digital Monitoring and Adaptive Oversight

Deep learning applications in healthcare demonstrate capacity for pattern detection in imaging and clinical signals (Esteva et al., 2021). In trials, AI-driven monitoring can detect adverse events across subpopulations in near real time, reducing harm and improving retention.

Regulatory frameworks from the FDA (2023) and EMA (2021) emphasize validation and data integrity in computerized systems. Alignment with these guidelines ensures that AI-enhanced oversight meets compliance standards.

Economic Implications

AI-enabled efficiencies may mitigate rising R&D costs (DiMasi et al., 2016) and key cost drivers in trials (Sertkaya et al., 2016). Improved recruitment speed and reduced failure rates (Fogel, 2018) enhance sustainability. By improving representativeness, AI may reduce costly post-marketing corrections.

DISCUSSION

The integration of AI/ML into RCTs represents a paradigm shift from reactive inclusion policies toward proactive, data-driven equity engineering. However, technology alone cannot dismantle structural inequities. Mortality disparities (Peterson et al., 2018) and chronic disease burdens (CDC, 2021) arise from social determinants beyond algorithmic reach. Therefore, AI systems must be embedded within broader policy reforms.

Algorithmic bias poses significant risks. Models trained on historically skewed datasets may perpetuate exclusion. Federated learning mitigates but does not eliminate this concern. Transparent auditing, inclusive data sourcing, and regulatory oversight are essential.

Digital divides also threaten equity. While EHR adoption is high (HealthIT.gov, 2023), disparities in digital literacy and access persist. Cloud computing solutions (Rajan et al., 2021) must account for infrastructural inequities across regions.

Future research should empirically test AI-driven recruitment models in multi-site trials, evaluate patient trust dynamics, and quantify cost-effectiveness.

CONCLUSION

Equity in randomized clinical trials is both a scientific imperative and a moral obligation. Persistent disparities in mortality, chronic disease, oncology genomics, and infectious disease outcomes underscore the urgency of reform. AI/ML technologies-when integrated with digital health infrastructure, federated collaboration, and robust regulatory governance-offer transformative potential. Yet their success depends on deliberate ethical design, bias mitigation, and structural awareness. The path forward requires interdisciplinary collaboration and empirical validation to ensure that the next generation of clinical trials reflects the diversity of the populations they aim to serve.

REFERENCES

1. Abbidi, S.R., Sinha, D. AI/ML-based strategies for enhancing equity, diversity, and inclusion in randomized clinical trials. *Trials* (2026). https://doi.org/10.1186/s13063-026-09537-2
2. Berchuck, J.E., et al. (2022). The prostate cancer androgen receptor cistrome in African American men associates with upregulation of lipid metabolism and immune response. Cancer Research, 82, 2848-2859.
3. Bhupathiraju, S.N., & Hu, F.B. (2016). Epidemiology of obesity and diabetes and their cardiovascular complications. Circulation Research, 118, 1723-1735.
4. Centers for Disease Control and Prevention. (2021). Diabetes: Advancing Health Equity.
5. DiMasi, J.A., Grabowski, H.G., & Hansen, R.W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics, 47, 20-33.
6. Esteva, A., et al. (2021). A guide to deep learning in healthcare. Nature Medicine, 25, 24-29.

7. Flores, L.E., et al. (2021). Assessment of the inclusion of racial/ethnic minority, female, and older individuals in vaccine clinical trials. JAMA Network Open, 4(2), e2037640.

8. Fogel, D.B. (2018). Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: a review. Contemporary Clinical Trials Communications, 11, 156-164.

9. HealthIT.gov. (2023). Office-based Physician Electronic Health Record Adoption.

10. Hirsch, J.S., et al. (2020). Acute kidney injury in patients hospitalized with COVID-19. Kidney International, 98, 209-218.

11. Kaitin, K.I. (2010). Deconstructing the drug development process: the new face of innovation. Clinical Pharmacology & Therapeutics, 87, 356-361.

12. Kershaw, K.N., & Pender, A.E. (2016). Racial/ethnic residential segregation, obesity, and diabetes mellitus. Current Diabetes Reports, 16, 108.

13. Miotto, R., Li, L., Kidd, B.A., & Dudley, J.T. (2016). Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. Scientific Reports, 6, 26094.

14. Peres, L.C., et al. (2022). Racial and ethnic differences in clonal hematopoiesis, tumor markers, and outcomes of patients with multiple myeloma. Blood Advances, 6, 3767-3778.

15. Peterson, K., Anderson, J., Boundy, E., Ferguson, L., McCleery, E., & Waldrip, K. (2018). Mortality disparities in racial/ethnic minority groups in the Veterans Health Administration: an evidence review and map. American Journal of Public Health, 108, e1-e11.

16. Rajan, R., Pillai, R., & John, A. (2021). Cloud computing in healthcare: Opportunities and challenges. Health Information Science and Systems, 9, 1-8.

17. Sertkaya, A., Wong, H.H., Jessup, A., & Beleche, T. (2016). Key cost drivers of pharmaceutical clinical trials in the United States. Clinical Trials, 13(2), 117-126.

18. Sheller, M.J., Edwards, B., Reina, G.A., Martin, J., Pati, S., & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. Scientific Reports, 10, 12598.

19. U.S. Food and Drug Administration. (2023). Digital Health Innovation Action Plan.

20. European Medicines Agency. (2021). Guideline on computerised systems and electronic data in clinical trials.

21. ClinicalTrials.gov. (2024). Trends, charts, and maps.

22. Centers for Disease Control and Prevention. (2021). Diabetes: Advancing Health Equity.