Available online on 15.06.2026 at <http://jddtonline.info>

Journal of Drug Delivery and Therapeutics

Open Access to Pharmaceutical and Medical Research

Copyright © 2026 The Author(s): This is an open-access article distributed under the terms of the CC BY-NC 4.0 which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited



Open Access Full Text Article



Review Article

Ethics of AI in Patient Recruitment and Diversity

Neeraj Singh Kholiya* , Abhijeet Ojha , Arun Kumar Singh , Vikas Bhatt

Faculty of Pharmaceutical Sciences (FPS), Amrapali University (Haldwani), India

Article Info:



Article History:

Received 26 April 2026
Reviewed 10 May 2026
Accepted 02 June 2026
Published 15 June 2026

Cite this article as:

Kholiya NS, Ojha A, Singh AK, Bhatt V, Ethics of AI in Patient Recruitment and Diversity, Journal of Drug Delivery and Therapeutics. 2026; 16(6):263-271 DOI: <https://doi.org/10.22270/jddt.v16i6.7815>

For Correspondence:

Neeraj Singh Kholiya, Faculty of Pharmaceutical Sciences (FPS), Amrapali University (Haldwani), India

Abstract

By 2026, global clinical research has seen a shift from voluntary diversity programs to mandatory regulation. This report examines the ethical and technical issues at the interface of artificial intelligence (AI) and demographic representation in trial enrollment. In the U.S. under the Food and Drug Administration's Diversity Action Plans and in India's New Drugs and Clinical Trials (NDCT) Rules of 2026, sponsors must now demonstrate that trial groups reflect the real-world diversity of the intended patient population.^{1, 2, 3} AI tools like TrialGPT and DocTr have emerged as key solutions, reporting up to 98.7% reduction in screening time and a 58% improvement in matching quality through analysis of large electronic health record (EHR) datasets.^{4, 5, 6} However, these technologies introduce significant ethical concerns, including "black box" epistemic opacity and algorithmic bias that can institutionalize healthcare inequalities.^{7, 8, 17} A persistent digital divide also excludes groups that lack high-speed connectivity or digital literacy.^{9, 10} This paper argues that although AI is central to achieving large-scale diversity in clinical research results, it must be implemented within governance frameworks such as the EU AI Act and WHO ethical guidelines, which guarantee human oversight, transparency, and data justice.^{10, 11, 12, 13} We recommend adherence to PRISMA 2020 reporting standards to ensure methodological rigor in the evaluation of AI-based recruitment interventions.^{12, 14}

Keywords: Ethics, Clinical Trials, Patient Recruitment, Diversity, Artificial Intelligence, Bioethical

1. Introduction

By 2026, the clinical research landscape has undergone a profound transformation from voluntary diversity initiatives to mandatory compliance. The FDA and CDSCO in India have both instituted strict measures requiring that clinical trials include representative populations from the diverse groups that medical products are intended to treat.^{1, 2, 3} Historically, clinical research has suffered from serious representation gaps; a 2019 analysis found that only 4% of racial and ethnic minorities were included in published randomized controlled trials.⁷ This persistent lack of representation is both a scientific and a social justice issue; it reflects a failure to capture genetic, sexual, age, and ethnic diversity, which in turn leads to the failure to identify adverse effects and poor drug responses in excluded groups.^{8, 17}

Artificial intelligence (AI) has emerged at the forefront of efforts to scale up recruitment and improve selection to fulfill these new regulatory requirements.^{4, 5, 6, 15} While these technologies may reduce human error and extend geographic reach, they also risk perpetuating historical biases through "black box" algorithms.^{7, 8, 22} This review examines the technical, ethical, and

regulatory dimensions of AI in patient recruitment, reporting on how the rise of intelligent software and simultaneously tighter regulation is transforming what we term health equity.^{10, 11, 16}

2. The Global Regulatory Mandate for Clinical Trial Diversity

2.1 United States: FDORA and Diversity Action Plans

A new era in US clinical research commenced with the Food and Drug Omnibus Reform Act of 2022 (FDORA), passed as part of the Consolidated Appropriations Act of 2023.³ FDORA added a requirement for Diversity Action Plans (DAPs) for Phase 3 and other pivotal studies of drugs, biologics, and medical devices requiring an investigational device exemption (IDE). DAPs must detail target subject numbers stratified by race, ethnicity, sex, and age, and must be submitted with Investigational New Drug (IND) applications and Phase 3 protocol submissions.^{1, 3}

The FDA issued draft guidance on DAPs in June 2024, outlining sponsor requirements for the form and manner of DAPs and their submission.¹ Notably, the statutory form and manner provisions of final DAP

guidance carry the force of law. The compliance obligation was set to activate 180 days after finalization, with final guidance statutorily due by June 2025. In January 2025, however, the FDA quietly removed its draft guidance from its website following executive orders against diversity, equity, and inclusion programs.^{1, 19}

This sudden removal created confusion for drug companies preparing Phase 3 trial submissions, since the statutory obligation requiring diversity in clinical trials under FDORA remains in effect even without implementing guidance.³ Despite the federal-level uncertainty, some institutions are independently advancing diversity recruitment. In cases where sponsors cannot meet planned enrollments, FDORA requires a report and mitigation action plan explaining the failure to comply, establishing that diversity is now tracked as an integral part of the drug development process.^{1, 3}

2.2 India: NDCT Amendment Rules 2026

India's regulatory landscape has been meaningfully reshaped by the Central Drugs Standard Control Organisation (CDSCO), under the Ministry of Health and Family Welfare, through the New Drugs and Clinical Trials (NDCT) Amendment Rules 2026.² These rules pursue a dual goal: to enhance the business climate for India's clinical research ecosystem and to harness the nation's remarkable diversity as a resource in scientific and regulatory decision-making. India's 1.4 billion inhabitants offer unmatched genetic, socioeconomic, linguistic, and cultural diversity a resource that can be ethically leveraged to enhance the external validity of international multiethnic clinical trials.^{2, 20}

Under the 2026 Amendment, Rule 52 introduces a system of "prior intimation" that dispenses with the need for a test license for specific low-risk drug-development activities and caps the time for regulatory approval from 90 to 45 days.² Under the revised rules, Ethics Committees are now required to structure the protection of vulnerable participants with written procedures. They must also oversee the compressed timelines that AI-powered trials tend to operate on, because an algorithm that identifies a patient in seconds still owes that patient adequate time to understand what they are being asked to consent to.^{2, 20, 21}

2.3 Bioethical Principles: Justice and the Belmont Report / CIOMS

The researchers who invest time and effort in clinical research should ensure that communities contributing to that research also benefit from it.¹³ In the context of clinical trial diversity, this principle extends beyond representational numbers; it requires that the results of trials, and the knowledge they produce, flow back to the communities that were included in them.^{13, 20}

The CIOMS International Ethical Guidelines push this further across borders, specifically addressing what is owed to women, children, and people in resource-limited settings when research is conducted in or about their communities.²⁰ Research conducted in low- and middle-income countries primarily for cheaper recruitment but generating knowledge useful mainly to wealthy populations is ethically exploitative, not neutral.²⁰ This is the moral standard against which AI recruitment tools must ultimately be judged: a tool that makes enrollment faster but leaves the distribution of benefit as unequal as before has not solved the equity problem it has merely automated it.^{8, 10, 21}

Table 1: PRISMA 2020 Item Relevance to AI Recruitment Studies

PRISMA 2020 Item	Relevance to AI Recruitment Studies	Reporting Consideration
Title / Abstract	Must specify AI tools used and populations targeted	State algorithm type (NLP, ML, LLM) and diversity metrics in abstract
Search Strategy	EHR, registry, and SDoH data sources are AI-specific study databases	Document data source, date range, and pre-processing steps
Eligibility Criteria	AI-defined criteria may differ from clinician-defined criteria	Report both algorithmic and clinical eligibility thresholds and any divergences
Risk of Bias	Training data bias, model validation population, fairness metrics applied	Report fairness metrics (e.g., equalized odds, demographic parity) by demographic subgroup
Results: Participants	Demographic breakdown of AI-identified vs. enrolled vs. excluded participants	Flow diagram disaggregated by race, ethnicity, sex, age, and SES
Discussion: Limitations	Algorithm opacity, digital divide, consent process limitations	Address explainability limitations and digital access barriers explicitly

Source: Adapted from Page et al. (2021)¹²

3. Technical Architectures of AI-Driven Recruitment

3.1 Sources and Modes of Data Processing

AI-based patient enrollment programs assemble large datasets that would be unmanageable without technological assistance.^{4, 5, 6} At the foundation are electronic health records (EHRs), which include structured clinical data fields (diagnoses, lab results, medications, vital signs) as well as unstructured clinical notes documenting clinical reasoning. Also incorporated are patient registries, claims databases, genomic data, and social determinants of health (SDoH) including stable housing, employment status, transport access, and health literacy to enable deeper eligibility evaluation.^{9, 15}

Natural language processing (NLP) is the key technology enabling access to unstructured clinical notes, which contain data such as performance status, historical treatments, and assessments of functional status not captured in structured fields.^{4, 22} Sophisticated NLP pipelines extract, harmonize, and reason over these notes, helping to make eligibility decisions closer to those made by a clinical research coordinator.^{4, 15}

3.2 Main Platforms: TrialGPT, DocTr, and Tempus

Among the most advanced AI trial recruitment systems, TrialGPT created in collaboration between the University of Illinois and the National Institutes of Health is a methodological landmark. TrialGPT is built on large language models using a zero-shot approach, achieving 87.3% accuracy in eligibility assessments during comparative testing.⁴ From an operational perspective, its speed is notable: the system processed more than 1,000 patients in two hours, a 98.7% increase in speed over manual verification.⁴ This is significant given that recruitment delays contribute an average of six to eight months to Phase 3 trial duration, and that 80% of clinical trials experience recruitment delays.^{15, 16}

The DocTr model employs deep learning architectures trained on claims data and unstructured clinical documents, with built-in fairness structures rather than post-hoc fairness adjustments.⁵ Head-to-head with manual methods, DocTr produced a 58% improvement in match quality. DocTr's genetic optimization sub-modules apply sample re-weighting to underrepresented demographic groups, pushing racial and ethnic fairness metrics up by as much as 25% compared to unmodified baselines an argument in design that fairness belongs in the architecture, not as an afterthought.^{5, 17}

Tempus' Patient Query tool integrates large language models and machine learning with clinical review to improve patient identification, achieving a 28.54% rise in identification rates compared to prior methods.⁶ Beyond identification, AI-powered recruiting tools increase the geographic and demographic breadth of recruitment by identifying potential participants in the community who are often missed by academic medical center-based pipelines an aspect with clear ethical implications for selection bias.^{6, 30}

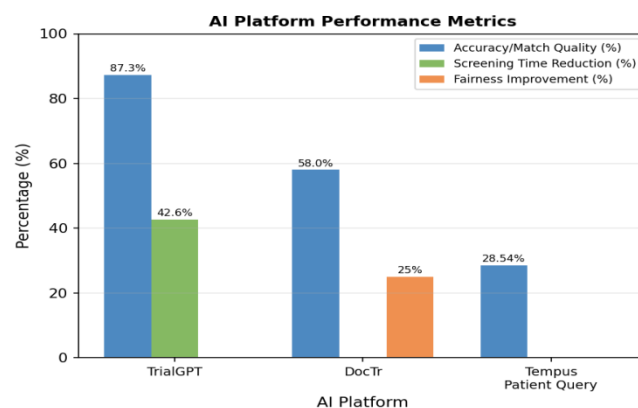


Figure 1. Performance metrics of AI-driven clinical trial recruitment platforms

Sources: Jin et al. (2024), Gao et al. (2021), Tempus (2024), Laredo et al. (2024), NIH All of Us (2024).

3.3 Federated Learning and Privacy-Preserving Architectures

A major ethical constraint on AI recruitment tools is balancing the need to train on diverse demographic data which requires pooling records across populations, hospitals, and communities with the obligation to protect patient privacy.^{27, 28} Federated learning addresses this by training models simultaneously across many datasets while keeping patient data within the local healthcare institution. In health care networks like TriNetX, demographic analyses are performed within each institution and only model weights are shared between institutions.^{26, 27} This approach enables large-scale demographic studies while retaining data within institutional control and minimizing privacy exposure.^{27, 28, 29}

4. Algorithmic Bias and the Risks of Unintentional Exclusion

4.1 Sources and Mechanisms of Bias

Label bias occurs when the outcomes on which AI models are trained are proxies for inequitable healthcare. The archetypal example is the healthcare algorithm examined by Obermeyer et al. (2019), which used healthcare costs as an indicator of health need.⁷ Since Black patients historically had lower costs than White patients with similar health needs due to historic barriers to care the algorithm underpredicted the health needs of Black patients and resulted in fewer resources being allocated to them. Trial populations recruited through such an approach would echo historic injustices.^{7, 8}

Algorithms trained on male/female binary data, or sourced from single-disease trial cohorts, may perform poorly for gender-diverse individuals or those with multiple comorbidities, producing inaccurate eligibility assessments or failing to identify these individuals entirely.^{8, 17, 24} Historical underrepresentation in EHR datasets further amplifies the potential for such systematic exclusion.^{9, 25}

4.2 The Epistemic Opacity Problem

Transformer-based large language models such as TrialGPT do not provide explanations for their outputs. They may include or exclude a patient from a priority list without indicating which elements of the data drove that decision information that clinicians, ethicists, and patient advocates need to evaluate.^{4, 22} This epistemic void has negative consequences across multiple domains of clinical research ethics: it impedes truly informed consent, and in the event of post-market scrutiny for recruitment bias, tracing algorithmic decision pathways is exceptionally difficult without access to the decision logic itself.^{7, 22}

The EU AI Act's designation of AI systems in healthcare as high-risk, with requirements for explainability, human oversight, and auditing, is a direct regulatory response to this opacity problem.¹¹ However, the challenge of producing genuine explanations from complex neural networks rather than plausible post-hoc rationalizations remains significant, and the sufficiency of current explainability methods to demonstrate true ethical and regulatory transparency is actively debated.^{22, 23}

4.3 Mitigation Approaches

Various technical methods have been proposed for detecting and mitigating AI bias. Pre-processing approaches include balanced class distribution via class re-sampling, re-weighting in favor of minority class examples, and data augmentation including synthetic data generation to increase representation of small minority groups as demonstrated in DocTr's genetic optimization sub-modules, which report 25% improved fairness metrics over baseline models.^{5, 17, 24}

In-processing methods embed fairness constraints within the training objective itself, simultaneously minimizing predictive loss and some degree of statistical inequality. Adversarial debiasing where a second model is trained to predict demographic group from the primary model's outputs and is penalized for accuracy in doing so is a prominent example of this approach.^{17, 24, 25} Post-processing methods operate on a trained model's predictions by adjusting thresholds to equalize false positive or false negative rates across demographic groups.^{17, 18}

All these approaches are subject to a fundamental conceptual limitation: the multiple definitions of algorithmic fairness demographic parity, equalized odds, calibration, and individual fairness are mutually inconsistent in most real-world situations.^{18, 24, 25} Satisfying one fairness criterion will generally mean that other criteria are violated. This is not merely an algorithmic challenge but an ethical one, requiring explicit deliberation about which conception of fairness is most appropriate for a given trial context, and why.^{18, 24}

5. Socio-Technical Disparities and the Digital Divide

Decentralized clinical trials (DCTs), which use remote monitoring, wearables, and digital health technologies to increase geographic diversity, may paradoxically widen equity gaps if they include only digitally capable participants, creating trial populations that are more digitally engaged but not necessarily more demographically representative.^{30, 31}

In India, where the 2026 NDCT Rules aim to harness the diversity of the Indian population for clinical trials, large infrastructure gaps remain.² Rural India, which comprises a substantial proportion of the population, frequently lacks access to diagnostic resources such as MRI and ultrasound, as well as the broadband connectivity that many digital trial technologies assume.^{2, 32} AI-based recruitment methods that presuppose digital connectivity are likely to exclude precisely the populations that the NDCT Rules are designed to include.^{2, 9}

"Digital bridge" programs in which algorithm-based outreach is complemented and delivered by human community engagement have been proposed to close this gap.¹⁴ The NIH's All of Us program, which reported enrolling over 80% of underrepresented groups into the study, demonstrates that AI and human interaction in recruitment can be complementary rather than substitutional.¹⁴ Culturally relevant community outreach and comprehensive language support for participants integrating both digital and analog aspects of research present a more equitable model of AI-augmented recruitment.^{14, 32, 33}

Ethics committees have a role to play in assessing the digital literacy requirements of trial participants as they review AI-based recruitment protocols.^{20, 21} Ethics committees must create standard operating procedures for vulnerable populations in the context of accelerated AI-mediated trial procedures, ensuring that the rate at which AI presents and processes recruitment information does not outpace the patient's capacity to understand and make autonomous decisions.^{2, 20, 21}

6. Informed Consent in the Age of Algorithmic Recruitment

Informed consent is a fundamental element of human subjects research, and AI-based recruitment complicates its implementation in meaningful ways.^{13, 34} The traditional model of informed consent requires that a patient receive comprehensible information about the research its purpose, procedures, alternatives, and risks from a qualified clinical expert, leading to a free and uninfluenced decision. AI in recruitment modifies this model at multiple points.^{13, 21}

When AI is used to identify patients for inclusion in a study, those patients may have little knowledge of the algorithmic role played in their selection, which data were used, and which criteria the algorithm applied. Consent processes should explain not only trial details

but also the recruitment process itself what data the algorithm used (for instance, medical history, genetic information, or behavioral data), how that data was used to determine eligibility, and what options a participant has if they consider the recruitment process unfair or inaccurate.^{21, 34, 35}

A significant power asymmetry exists between patients some of whom may have serious illness and few treatment alternatives and the research institutions enrolling them. The appearance of access to potentially beneficial interventions provided by AI outreach can make participation feel obligatory, particularly if patients believe it to be their only route to needed care.^{13, 21} The knowledge gap between algorithm developers and participants further compounds this asymmetry.^{8, 22}

Data consent to use an individual's health information for AI model training is frequently obtained under very general institutional consent terms that participants may not fully understand.^{21, 34} The use of health data to train algorithms that will select future trial participants may challenge the adequacy of the initial consent to collect that data. Models of dynamic consent, in which patients are able to revisit and modify their data-use preferences over time via digital platforms, represent a promising alternative, though they require substantial infrastructure and participant engagement for success.^{21, 35}

The WHO's bioethical principle of Autonomy requires the development of consent processes that genuinely include the patient beyond checkbox exercises.¹⁰ For AI in recruitment, this means designing consent that clearly explains what the algorithm is doing in terms accessible to patients of varying health and digital literacy, allows for discussion and explanation, and provides patients the ability to withdraw from data use at any time without affecting their access to care.^{10, 21, 34}

7. Governance Frameworks and the Path to Ethical AI Recruitment

7.1 The EU AI Act

The EU Artificial Intelligence Act, which came into effect in 2024, represents to date the most comprehensive regulatory framework for AI governance, with direct implications for AI in health and clinical research.¹¹ Under the Act's risk taxonomy, AI in health and clinical care including clinical trial recruitment is classified as high-risk. High-risk AI systems are subject to pre-market evaluation requirements, technical documentation of training data and design decisions, logging of decisions to support retrospective audit, human-in-the-loop components enabling review and override of algorithmic outputs, and post-marketing performance surveillance.¹¹

These obligations directly address the core ethics of AI recruitment. Logging requirements create a traceable record of enrollment decisions that auditors can examine when bias is suspected after the fact.^{11, 22} Transparency requirements give informed consent real substance, moving it toward something patients can genuinely act upon. Human oversight requirements eliminate the epistemic opacity problem by ensuring no algorithmic outcome is final without human review. The Act's prohibition on AI systems that manipulate vulnerable people's characteristics to alter their behavior may also apply to coercive algorithmic outreach directed at individuals with life-threatening illness.^{11, 21}

7.2 WHO Bioethical Principles for AI Recruitment

The World Health Organization's global guidance on ethics and governance of AI for health (2021) applies the classical bioethical principles of Beneficence, Non-Maleficence, Autonomy, and Justice to the use of AI in health.¹⁰ This framework provides the evaluative structure for assessing AI recruitment tools against broader ethical obligations to patients and communities.^{10, 13}

Table 2. WHO Bioethical Principles Applied to AI Patient Recruitment

Principle	Requirement in AI Recruitment	Current Challenges	Mitigation Strategies
Beneficence	AI should genuinely improve diversity and access, not merely increase speed	Efficiency gains may not translate to equity gains if bias is embedded	Mandate fairness metrics alongside accuracy metrics in validation studies
Non-Maleficence	AI must not cause harm through exclusion, privacy breach, or coercive outreach	Algorithmic exclusion is invisible and difficult to attribute	Require algorithmic audit trails; prohibit fully automated exclusion decisions
Autonomy	Patients must meaningfully understand and consent to algorithmic involvement	Black-box systems undermine informed consent; power asymmetry undermines voluntariness	Develop AI-specific consent language; implement dynamic consent platforms
Justice	Benefits and burdens of research must be equitably distributed	Training data bias risks reproducing historical exclusions at algorithmic scale	Require diverse training datasets; mandate subgroup fairness reporting in trials

Source: Adapted from WHO (2021)¹⁰

7.3 Proposed Governance Recommendations

Based on the regulatory frameworks and ethical analysis presented above, the following governance recommendations are proposed for AI patient recruitment:

Mandatory algorithmic fairness audits: All AI recruitment systems deployed as part of FDORA DAPs or equivalent national diversity programs should undergo pre-deployment audit for standard fairness metrics including demographic parity, equalized odds, and calibration disaggregated by race, ethnicity, sex, age, and where available socioeconomic status. Reports should be provided to sponsors, ethics boards, and regulators.^{1, 3, 11, 17, 18}

Human-in-the-loop requirements: No AI system should make any recruitment decision inclusion, exclusion, or eligibility classification without human oversight. Algorithms should produce ranked candidate lists and eligibility rationales for review, override, and documentation by qualified clinical staff.^{11, 21, 22}

Audit trail and explainability requirements: AI-based recruitment systems should maintain complete audit trails of all eligibility decisions, together with the data used and algorithmically generated rationale, for

the period following trial closure, with tamper-proof records. Where the algorithm is not fully explicable, this should be disclosed to ethics committees and trial participants.^{11, 22, 23}

Diversity standards for training data: Data used to train AI recruitment tools should meet minimum diversity criteria, with reporting of data used and its known limitations. Where data for minority groups are sparse, sponsors should disclose the expected effect on model generalization for those groups.^{8, 17, 25}

Intersectional equity in digital access: Framing universal digital access as a baseline assumption is not a methodologically neutral choice it implicitly excludes groups without digital connectivity. Trial designs should incorporate analog outreach pathways and digital bridge programs to ensure that populations most needed for diversity are genuinely reachable.^{9, 14, 30, 32}

Third-party auditing: Clinical data has never been trusted to self-audit. An equivalent infrastructure of independent, third-party auditing should be extended to AI recruitment systems, with mandates covering both bias detection and regulatory compliance. Sponsor self-reporting is not a credible substitute when the question is which patients were found and which were never sought.^{11, 17, 21}

Table 3. Comparative Performance of AI-Driven Recruitment Platforms

Platform	Technology / Approach	Accuracy & Performance Metrics	Efficiency Gains	Fairness & Equity Features	Limitations & Concerns
TrialGPT	Large language model (zero-shot, GPT-based)	87.3% criterion-level accuracy (vs. expert 88.7–90.0%); recall >90% of relevant trials	42.6% reduction in screening time; outperforms best baselines by 43.8% in ranking	Faithful eligibility explanations; human-interpretable rationales per criterion	Epistemic opacity in LLM decision-making; no built-in demographic fairness constraints; validated on synthetic patients
DocTr	Deep learning on claims data and unstructured clinical documents	58% improvement in match quality over manual methods	Built-in fairness architecture; genetic optimisation sub-modules	25% improved racial/ethnic fairness metrics vs. unmodified baseline	Trained on claims data that may embed historical cost-based biases; fairness–accuracy trade-off when constraints are tightened
Tempus Patient Query	LLM + ML integration with clinical review; oncology-focused real-world data	28.54% improvement in patient identification vs. prior methods	Extends geographic and demographic recruitment reach beyond academic medical centres	No dedicated fairness sub-modules reported; relies on clinical review layer for bias oversight	Primarily validated in oncology; limited transparency for non-oncology applications; presupposes EHR connectivity

Source: Adapted from Jin et al. (2024), Gao et al. (2021), Tempus (2024), Gichoya et al. (2025).

Table 4. Summary of Research Gaps, Evidence Quality, and Priority Directions for AI in Clinical Trial Recruitment

Research Gap Area	Current Evidence Status	Priority Level	Recommended Next Step
LMIC Applicability of AI Tools	Largely uncharted; most validation from US, Western Europe, and high-income East Asia	High	Prospective multi-country validation studies with India, Sub-Saharan Africa, and Latin America cohorts
Intersectional Fairness Measurement	Chronically under-reported; most studies treat race, sex, and age as independent variables	High	Algorithm validation studies with pre-specified intersectional subgroup analysis (race × sex × SES)
Post-Recruitment AI Monitoring Ethics	Ethical frameworks for AI-powered attrition forecasting and adherence monitoring are largely absent	Moderate-High	Expert consensus framework development; qualitative studies with trial participants on surveillance acceptability
Real-World Diversity Outcomes of AI Recruitment	Efficiency gains demonstrated; demographic diversity gains vs. AI-naïve trials not yet robustly established	High	Randomised or quasi-randomised comparisons of diversity endpoints in AI vs. non-AI recruitment arms

Source: Adapted from Laredo et al. (2024), Abbidi & Sinha (2026), Gichoya et al. (2023), Mathur et al. (2024).

8. Gaps in Research and Future Directions

8.1 Low- and Middle-Income Country and Global Contexts

Most current research on AI recruitment tools has been conducted within the United States, Western Europe, and high-income East Asian regions.^{16, 36} The applicability of these tools in low- and middle-income country (LMIC) settings such as India, Sub-Saharan Africa, and Latin America where health system infrastructure, EHR implementation, and data standards differ substantially from the settings in which algorithms were trained and validated, remains largely uncharted.^{9, 16} India's NDCT 2026 Rules present an opportunity to address this knowledge gap, but require directed research and regulatory investment.²

8.2 Intersectional Identity and Compound Marginalization

Many algorithmic fairness studies treat potential sources of bias — race, sex, age, socioeconomic status — such as independent variables.^{17, 24} In practice, individuals occupy multiple social categories simultaneously, and the effects of marginalization are often multiplicative for those belonging to multiple stigmatized groups.^{24, 25, 33} Intersectional fairness measuring model performance across intersections of multiple demographic groups is technically complex and chronically underreported in validation studies. Tools and criteria for evaluating intersectional bias represent a high-priority direction for future research.^{17, 18, 24}

8.3 Post-Recruitment Monitoring and Retention Ethics

The ethics literature on AI in clinical trials has focused predominantly on recruitment. However, AI technologies are increasingly used to forecast and minimize participant attrition, monitor adherence, and detect protocol violations raising distinct ethical concerns regarding surveillance, potential manipulation, and the collection of sensitive behavioral data.^{15, 16} Dedicated ethical frameworks for AI-powered trial monitoring, as distinct from recruitment, have yet to be developed at any comparable depth.^{21, 35}

8.4 Effectiveness of AI Diversity Strategies

While integrating AI into trial processes may increase efficiency, there is a lack of robust evidence that AI tools actually increase demographic diversity in trial populations, as opposed to increasing the efficiency of recruiting individuals already within established institutional pipelines.^{15, 16} Randomized or quasi-randomized comparisons of diversity outcomes between trials using AI recruitment methods and those that do not with pre-specified diversity endpoints are necessary to move beyond case studies and claims toward genuine evidence.^{12, 15, 23}

9. Conclusion

The use of AI in patient recruitment for clinical trials represents one of the most significant advances in recent clinical research. AI-driven platforms can screen for eligibility at a scale and pace that human researchers cannot match, and in doing so offer the genuine potential to broaden the geographic, demographic, and ethnic diversity of clinical trial populations beyond the

traditional academic medical center pipelines.^{4, 5, 6} In a regulatory environment that increasingly treats demographic representativeness as a condition of market approval, AI's capacity to reach patients at scale is not a competitive advantage it is a compliance necessity.^{1, 2, 3}

However, this ethical analysis makes clear that efficiency and justice are distinct properties, and that one can be achieved at the expense of the other if AI is not used responsibly. An algorithm trained on historically skewed data will efficiently find historically skewed patients at scale.^{7, 8} A system that excludes without explanation does not merely underperform it violates the transparency principles on which ethical research is founded.^{22, 23} And a recruitment strategy designed around assumptions of broadband access and digital fluency will, almost by design, screen out the very populations that diversity mandates exist to bring in.^{9, 30, 32}

The regulatory landscape is shifting, but unevenly. In the United States, the FDA's statutory obligation to require Diversity Action Plans under FDORA remains law, while the implementing guidance has disappeared from public view.^{1, 3} India's NDCT 2026 Amendment Rules open genuine possibilities in one of the world's most demographically rich patient populations.² The EU AI Act establishes a serious governance framework for high-risk AI systems including clinical trial recruitment.¹¹ Taken together with the WHO's bioethical principles, the Belmont Report, and the CIOMS Guidelines, these frameworks offer a foundation for responsible AI recruitment.^{10, 13, 20} But frameworks do not implement themselves. That work requires researchers, clinicians, ethicists, patient advocates, and regulators working in concert rather than in institutional silos.^{21, 33}

Priority research investments should address intersectional fairness methods, strategies effective in low- and middle-income country settings, the ethics of post-recruitment participant engagement, and rigorous testing of whether AI recruitment tools actually translate into improved real-world diversity outcomes rather than more efficient attainment of enrollment numbers that satisfy a DAP target while leaving the structural conditions producing health disparities untouched.^{9, 16, 17, 24, 33}

The aim is not a more efficient perpetuation of inequalities at machine pace, but a more efficient achievement of justice at algorithmic pace a just distribution of the benefits and burdens of clinical research, which AI, when ethically deployed and rigorously governed, can help to bring about.^{10, 13}

Acknowledgement: The authors would like to express their sincere gratitude to the faculty members for their valuable support and guidance throughout the preparation of this manuscript. These authors also acknowledge the efforts of all co-authors in successfully completing this work.

Author Contribution: Neeraj Singh Kholiya contributed to the literature review and drafting of the manuscript. Dr. Abhijeet Ojha, Dr. Arun Kumar Singh, and Vikas Bhatt contributed to concept development, guidance, review, and plagiarism checking of the manuscript. All authors have read and approved the final version of the manuscript.

Conflict of Interest: The authors declare no conflict of interest.

Ethical Approval: Not applicable.

References

1. U.S. Food and Drug Administration. Diversity Action Plans to Improve Enrollment of Participants from Underrepresented Populations in Clinical Studies. Draft Guidance for Industry. June 2024. Federal Register Notice 2024-14284.
2. Ministry of Health and Family Welfare, Government of India. New Drugs and Clinical Trials (Amendment) Rules, 2026. Gazette of India Extraordinary. 2026.
3. U.S. Congress. Food and Drug Omnibus Reform Act of 2022 (FDORA). Consolidated Appropriations Act, 2023. Pub. L. 117-328, Section 3601-3610.
4. Jin Q, Wang Z, Floudas CS, et al. Matching patients to clinical trials with large language models. *Nature Communications*. 2024;15:9074. DOI: <https://doi.org/10.1038/s41467-024-53081-z>
5. Gao S, Alawad M, Young MT, et al. DocTr: Document Transformer for Structured Information Extraction in Documents. *Proceedings of EMNLP*. 2021.
6. Tempus. Patient Query: AI-Powered Patient Identification for Clinical Trials. Tempus White Paper. 2024.
7. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447-453. DOI: 10.1126/science.aax2342
8. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M. Ethical machine learning in healthcare. *Annual Review of Biomedical Data Science*. 2021;4:123-144. DOI: <https://doi.org/10.1146/annurev-biodatasci-092820-114757>
9. Cantor MN, Thorpe L. Integrating data on social determinants of health into electronic health records. *Health Affairs*. 2018;37(4):585-590. DOI: <https://doi.org/10.1377/hlthaff.2017.1252>
10. World Health Organization. Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva: WHO; 2021.
11. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). Official Journal of the European Union. 2024. ELI: <http://data.europa.eu/eli/reg/2024/1689/oj>
12. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. DOI: <https://doi.org/10.1136/bmj.n71>
13. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research. Washington, DC: DHEW; 1979.
14. National Institutes of Health. All of Us Research Program.
15. Laredo SH, Nakagawa S, Shen MJ, et al. Artificial intelligence for optimizing recruitment and retention in clinical trials: a scoping review. *JAMIA*. 2024;31(11):2749-2762. DOI: <https://doi.org/10.1093/jamia/ocae205>

16. Abbidi SR, Sinha D. AI/ML-based strategies for enhancing equity, diversity, and inclusion in randomized clinical trials. *Trials*. 2026. DOI: <https://doi.org/10.1186/s13063-026-09537-2>
17. Gichoya JW, Thomas K, Celi LA, et al. Bias recognition and mitigation strategies in artificial intelligence healthcare applications. *npj Digital Medicine*. 2025;8(1):137. DOI: <https://doi.org/10.1038/s41746-025-01137-5>
18. Chen IY, Joshi S, Ghassemi M, Ranganath R. Ethical limitations of algorithmic fairness solutions in health care machine learning. *The Lancet Digital Health*. 2020;2(5):e221-e223. DOI: [https://doi.org/10.1016/S2589-7500\(20\)30065-0](https://doi.org/10.1016/S2589-7500(20)30065-0)
19. U.S. Food and Drug Administration. Enhancing the Diversity of Clinical Trial Populations Eligibility Criteria, Enrollment Practices, and Trial Designs. Guidance for Industry. November 2020.
20. Council for International Organizations of Medical Sciences (CIOMS). International Ethical Guidelines for Health-related Research Involving Humans. 4th ed. Geneva: CIOMS; 2016. DOI: <https://doi.org/10.56759/rgxl7405>
21. Mathur P, Ramirez AH, Rader DJ, et al. Ethical considerations in the design and conduct of clinical trials of artificial intelligence. *JAMA Network Open*. 2024;7(9):e2433952. DOI: <https://doi.org/10.1001/jamanetworkopen.2024.33952>
22. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. *arXiv*. 2017;1702.08608. DOI: <https://doi.org/10.48550/arXiv.1702.08608>
23. Park Y, Jackson GP, Foreman MA, Grion D, Hu J, Topaloglu U. Evaluating artificial intelligence in medicine: phases of clinical research. *npj Digital Medicine*. 2020;3:152. DOI: <https://doi.org/10.1038/s41746-020-00344-w>
24. Gichoya JW, McCoy LG, Celi LA, Ghassemi M. Algorithm fairness in artificial intelligence for medicine and healthcare. *Nature Biomedical Engineering*. 2023;7(6):719-742. DOI: <https://doi.org/10.1038/s41551-023-01056-8>
25. Xu J, Xiao Y, Wang WH, et al. Algorithmic fairness in computational medicine. *eBioMedicine*. 2022;84:104250. DOI: <https://doi.org/10.1016/j.ebiom.2022.104250>
26. Zhang FD, Shuai Z, Kuang K, Wu F, Zhuang Y, Xiao J. Unified fair federated learning for digital healthcare. *Patterns*. 2023;5(1):100907. DOI: <https://doi.org/10.1016/j.patter.2023.100907>
27. Sadilek A, Liu L, Nguyen D, et al. Privacy-first health research with federated learning. *npj Digital Medicine*. 2021;4(1):132. DOI: <https://doi.org/10.1038/s41746-021-00489-2>
28. Topaloglu MY, Morrell EM, Rajendran S, Topaloglu U. In the pursuit of privacy: the promises and predicaments of federated learning in healthcare. *Frontiers in Artificial Intelligence*. 2021;4:746497. DOI: <https://doi.org/10.3389/frai.2021.746497>
29. Goodson N, Wicks P, Morgan J, Hashem L, Callinan S, Reites J. Opportunities and counterintuitive challenges for decentralized clinical trials to broaden participant inclusion. *npj Digital Medicine*. 2022;5(1):58. DOI: <https://doi.org/10.1038/s41746-022-00603-y>
30. Aiyegbusi OL, Davies EH, Myles P, et al. Digitally enabled decentralised research: opportunities to improve the efficiency of clinical trials and observational studies. *BMJ Evidence-Based Medicine*. 2023;28(5):306-312. DOI: <https://doi.org/10.1136/bmjebm-2023-112253>
31. Mayo Clinic Proceedings. Health disparities, clinical trials, and the digital divide. *Mayo Clinic Proceedings*. 2023;98(12):1875-1887. DOI: <https://doi.org/10.1016/j.mayocp.2023.05.003>
32. Waters M. AI meets informed consent: a new era for clinical trial communication. *JNCI Cancer Spectrum*. 2025;9(2):pkaf028. DOI: <https://doi.org/10.1093/jncics/pkaf028>
33. Wang T, Emami E, Jafarpour D, Tolentino R, Gore G, et al. Integrating equity, diversity, and inclusion throughout the lifecycle of artificial intelligence for healthcare: a scoping review. *PLOS Digital Health*. 2025;4(7):e0000941. DOI: <https://doi.org/10.1371/journal.pdig.0000941>
34. Goodson N, Wicks P, Morgan J, Hashem L, Callinan S, Reites J. Opportunities and counterintuitive challenges for decentralized clinical trials to broaden participant inclusion. *npj Digital Medicine*. 2022;5(1):58. DOI: <https://doi.org/10.1038/s41746-022-00603-y>
35. Wang T, Emami E, Jafarpour D, Tolentino R, Gore G, et al. Integrating equity, diversity, and inclusion throughout the lifecycle of artificial intelligence for healthcare: a scoping review. *PLOS Digital Health*. 2025;4(7):e0000941. DOI: <https://doi.org/10.1371/journal.pdig.0000941>
36. Xu J, Xiao Y, Wang WH, et al. Algorithmic fairness in computational medicine. *eBioMedicine*. 2022;84:104250. DOI: <https://doi.org/10.1016/j.ebiom.2022.104250>