

Risk Management in AI-Assisted Clinical Trials: Regulatory and Ethical Dimensions

Filip Radonić, Singidunum University

Abstract – The integration of artificial intelligence (AI) models into clinical trials has accelerated significantly over the past decade, offering substantial potential to optimise patient recruitment, endpoint analysis, and pharmacovigilance. However, this integration also introduces a complex web of regulatory and ethical risks that must be systematically managed. This paper examined the principal risk categories associated with AI deployment in clinical research, with particular focus on algorithmic bias, the erosion of informed consent through opaque decision-making, and the evolving regulatory landscape shaped by the EU AI Act (Regulation EU 2024/1689), the European Medicines Agency (EMA) Reflection Paper on AI, and the U.S. Food and Drug Administration (FDA) draft guidance on AI in drug development. A structured review of peer-reviewed literature and regulatory documents was conducted. The findings indicated that, while regulatory convergence between the FDA and EMA is advancing, significant gaps persist in practice concerning transparency, explainability, and equitable participant selection. The paper proposed a multi-layered risk governance framework that integrates lifecycle monitoring, ethics review, and human oversight as foundational components of responsible AI use in clinical trials.

Keywords – Artificial intelligence; clinical trials; risk management; regulatory compliance; algorithmic bias; EU AI Act; informed consent; EMA; FDA

I. INTRODUCTION

Artificial intelligence (AI) models are transforming the design, execution, and monitoring of clinical trials. Sponsors now routinely employ machine learning (ML) algorithms for tasks ranging from patient recruitment and stratification to safety signal detection and adaptive trial design. The promise of AI in this domain is considerable: by processing large, heterogeneous

datasets at a speed and scale that exceeds human capacity, AI models can accelerate drug development timelines and reduce per-trial costs. Yet this promise is inseparable from a set of profound regulatory and ethical risks that, if unaddressed, may compromise patient safety, the integrity of scientific evidence, and public trust in clinical research.

This paper examines the regulatory and ethical risk landscape associated with AI deployment in clinical trials. It draws on recent regulatory guidance from the U.S. Food and Drug Administration (FDA)¹[1] and the European Medicines Agency (EMA),²[2] the EU Artificial Intelligence Act (AI Act),³[3] and peer-reviewed literature on algorithmic bias and informed consent. The central argument is that effective risk management in AI-assisted clinical trials requires not only compliance with formal regulatory requirements, but also a proactive governance culture that embeds transparency, accountability, and equity into every phase of the AI lifecycle.

II. REGULATORY LANDSCAPE

The regulatory environment governing AI in clinical trials has evolved rapidly since 2023. The FDA's 2023 Discussion Paper on AI in drug development, revised in February 2025, acknowledged the value of AI models while emphasising data transparency, algorithm explainability, and verifiable model performance.⁴[1] In January 2025, the FDA issued draft guidance titled "Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products," introducing a risk-based credibility assessment framework requiring sponsors to establish and document the credibility of any AI model used in regulatory submissions.⁵[16]

In parallel, the EMA published its Reflection Paper on AI in the Medicinal Product Lifecycle in September 2024, adopted by both the Committee for Human Medicinal Products (CHMP) and the Committee for Veterinary Medicinal Products (CVMP). The paper introduced the concepts of "high patient risk" and

¹U.S. Food and Drug Administration, "Using Artificial Intelligence and Machine Learning in the Development of Drug and Biological Products: Discussion Paper and Request for Feedback" (May 2023, revised February 2025). U.S. Department of Health and Human Services.

²European Medicines Agency, "Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle," EMA/CHMP/CVMP, adopted September 2024. Amsterdam: EMA.

³European Parliament and Council of the European Union, Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial

intelligence (Artificial Intelligence Act). Official Journal of the European Union, 12 July 2024.

⁴U.S. Food and Drug Administration, "Using Artificial Intelligence and Machine Learning in the Development of Drug and Biological Products: Discussion Paper and Request for Feedback" (May 2023, revised February 2025). U.S. Department of Health and Human Services.

⁵FDLI, "Regulating the Use of AI in Drug Development: Legal Challenges and Compliance Strategies," Food and Drug Law Institute, July 2025. Available at: <https://www.fdpi.org/2025/07/regulating-the-use-of-ai-in-drug-development/>.

"high regulatory impact" to categorise AI applications by the severity of potential consequences, and mandated lifecycle governance, human oversight, and data integrity across all development stages.⁶[2]

A landmark development occurred in January 2026, when the EMA and FDA jointly published ten guiding principles for good AI practice in the medicines lifecycle.⁷[4] These non-binding principles emphasise a human-centric, risk-based approach, multidisciplinary expertise, and transparent model development. They represent the first concerted EU-US regulatory convergence on AI and are intended to underpin future binding guidance in both jurisdictions.

The EU AI Act (Regulation EU 2024/1689), which entered into force on 1 August 2024, constitutes the world's first comprehensive legal framework for AI. It adopts a risk-tiered approach, classifying AI systems into prohibited, high-risk, transparency-requiring, and general-purpose categories. AI systems used in clinical trials, including patient recruitment tools, clinical decision support software, endpoint adjudication systems, and synthetic control arm generators, are highly likely to be classified as high-risk under Annex III of the Act.⁸[8] High-risk classification triggers extensive compliance obligations: comprehensive technical documentation, risk management plans, data governance aligned with the GDPR, transparency to deployers, human oversight mechanisms, and registration with EU competent authorities.⁹[9]

Importantly, the AI Act intersects with, but does not replace, existing pharmaceutical and medical device regulations, including the EU Clinical Trials Regulation (No 536/2014), the Medical Device Regulation (EU 2017/745), and the GDPR. Navigating this overlapping regulatory framework represents a significant compliance challenge for trial sponsors, particularly those operating across multiple jurisdictions.¹⁰[3]

III. ALGORITHMIC BIAS AS AN ETHICAL AND SCIENTIFIC RISK

Among the ethical risks posed by AI in clinical trials, algorithmic bias is perhaps the most pervasive and consequential. Bias in medical AI arises and compounds throughout the AI lifecycle: in data collection and labelling, in model architecture and training, in deployment settings, and in the interpretation of outputs.¹¹[6] The foundational principle that governs this risk is often described as "bias in, bias out",

which captures how biases embedded in training data propagate into model outputs and, ultimately, into clinical decisions.

In the context of clinical trials, algorithmic bias manifests most acutely in patient selection and recruitment. AI-driven tools that analyse electronic health records (EHRs) or genomic databases may systematically under-recruit patients from minority ethnic groups, elderly populations, or those with lower socioeconomic status, either because these groups are underrepresented in training data, or because the model's optimization objective inadvertently encodes existing healthcare inequities.¹²[7] When such biased tools are deployed at scale, they threaten the generalisability of trial findings and may expose already-marginalised populations to the dual harm of exclusion from trials and subsequent receipt of treatments validated on non-representative cohorts.

A widely cited demonstration of this risk was provided by Obermeyer et al. (2019), who showed that a commercially deployed healthcare algorithm assigned equal risk scores to Black and white patients despite Black patients being significantly sicker, because the model used healthcare costs as a proxy for medical need a variable that reflects historical patterns of racial discrimination in healthcare spending.¹³[7] This finding illustrated that algorithmic bias can be implicit and structural rather than deliberate, making it especially difficult to detect without systematic fairness auditing.

Traditional ethical frameworks for clinical research, including the Nuremberg Code, the Declaration of Helsinki, the Belmont Report, and the CIOMS Guidelines, enumerate fair participant selection as a core principle. A qualitative study by Youssef et al. (2024), drawing on interviews with investigators in AI-based diabetic retinopathy screening trials, found that ensuring equitable participant selection was one of the most challenging ethical issues unique to AI clinical trials.¹⁴[5] Traditional ethics review processes, designed for conventional research, may be ill-equipped to evaluate the fairness properties of ML models without specialist input from data scientists and ethicists.

Mitigation strategies include diverse dataset curation, algorithmic fairness audits prior to deployment, stratified performance evaluation across demographic subgroups, and continuous post-deployment monitoring. Both the EMA Reflection Paper and the FDA credibility assessment framework call for ongoing performance monitoring

⁶European Medicines Agency, "Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle," EMA/CHMP/CVMP, adopted September 2024. Amsterdam: EMA.

⁷European Medicines Agency and U.S. Food and Drug Administration, "Ten Principles for Good Artificial Intelligence Practice in the Medicines Lifecycle," Joint EMA-FDA Statement, January 2026.

⁸Busch, F., Kather, J.N., and Johner, C., "Navigating the EU AI Act: Implications for Regulated Digital Medical Products," *npj Digital Medicine* 7 (2024): 210. doi:10.1038/s41746-024-01213-6.

⁹Advarra, "Understanding the Impact of the New EU Artificial Intelligence Act on Clinical Research," Advarra Blog, February 2025. Available at: <https://www.advarra.com/blog/understanding-the-impact-of-the-new-eu-artificial-intelligence-act-on-clinical-research/>.

¹⁰European Parliament and Council of the European Union, Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial

intelligence (Artificial Intelligence Act). Official Journal of the European Union, 12 July 2024.

¹¹Cross, J.L., Choma, M.A., and Onofrey, J.A., "Bias in Medical AI: Implications for Clinical Decision-Making," *PLOS Digital Health* 3, no. 11 (2024): e0000651. doi:10.1371/journal.pdig.0000651.

¹²Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S., "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," *Science* 366, no. 6464 (2019): 447-453. doi:10.1126/science.aax2342.

¹³Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S., "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," *Science* 366, no. 6464 (2019): 447-453. doi:10.1126/science.aax2342.

¹⁴Youssef, A., et al., "Ethical Considerations in the Design and Conduct of Clinical Trials of Artificial Intelligence," *JAMA Network Open* 7, no. 9 (2024): e2432643. doi:10.1001/jamanetworkopen.2024.32643.

throughout the model lifecycle.¹⁵[2][14] The AI Act further reinforces this by mandating incident reporting and post-market surveillance for high-risk AI systems.¹⁷[3]

IV. INFORMED CONSENT AND THE EXPLAINABILITY PROBLEM

A second major ethical risk concerns the compatibility of AI-based decision-making with patients' right to informed consent. In clinical research, informed consent requires that participants understand the nature, purpose, risks, and foreseeable consequences of their participation. When AI models are involved in decisions about participant selection, treatment allocation, or safety monitoring, the question arises: can meaningful consent be obtained for decisions made, at least in part, by algorithms whose internal reasoning is not transparent?¹⁸[13]

The "black box" problem refers to the inability of many high-performing AI models, particularly deep neural networks, to provide human-interpretable explanations for their outputs. Chan (2023) argued that physicians' use of black-box AI undermines their ethical duty to advise patients, because communicating an AI recommendation without being able to explain the reasoning behind it erodes patient understanding, autonomy, and dignity.¹⁹[11] This argument has particular force in clinical trials, where participants occupy a vulnerable position and where the research-clinical boundary requires heightened ethical scrutiny.

The legal dimension of this problem is equally complex. Under European human rights instruments and patient rights law, the extent to which informed consent requires patient understanding of the AI's decision-making process remains contested. Cohen and Slottje (2024) concluded that the introduction of AI creates genuine uncertainty about disclosure obligations, particularly when AI systems generate recommendations that deviate from standard clinical guidelines without the physician being able to explain why.²⁰ [13]

Amann et al. (2020) provided a multidisciplinary framework for AI explainability in healthcare, distinguishing between interpretability (understanding the model's general structure)

and explainability (understanding why a specific output was produced for a specific patient).²¹[12] They argued that different stakeholders, including regulators, clinicians, and patients, require different levels and types of explanation, and that a one-size-fits-all approach is insufficient. The AI Act's transparency requirements for high-risk systems, which include clear instructions for use and disclosure of system limitations, provide a regulatory floor for explainability.²²[9] The GDPR's provisions on automated decision-making (Article 22) further require that data subjects be informed about the logic of automated decisions significantly affecting them and retain the right to obtain human review.²³[15]

V. MULTI-LAYERED RISK MANAGEMENT FRAMEWORK

Drawing on the regulatory and ethical analysis above, this paper proposes a multi-layered risk management framework for AI in clinical trials, structured around four interdependent pillars.

A. Pre-Deployment Risk Assessment

Before any AI model is deployed in a clinical trial, a structured credibility assessment should be conducted, aligned with the FDA's draft guidance and the EMA's Reflection Paper. This assessment should document: (1) the model's intended context of use; (2) the training, validation, and test datasets, including an explicit analysis of demographic representativeness; (3) performance metrics disaggregated by relevant subgroups; (4) potential failure modes and their patient safety consequences; and (5) the rationale for the chosen model complexity and any explainability trade-offs.²⁴²⁵

B. Ethics Review and Consent Redesign

Institutional review boards (IRBs) and ethics committees should be equipped with multidisciplinary expertise to evaluate AI-specific ethical risks, including fairness audits and explainability assessments. Informed consent documents for AI-assisted trials should clearly disclose the role of AI in participant selection, the nature and limitations of algorithmic

¹⁵European Medicines Agency, "Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle," EMA/CHMP/CVMP, adopted September 2024. Amsterdam: EMA.

¹⁶Chin, M.H., et al., "Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Health Care," JAMA Network Open 6, no. 12 (2023): e2345050. doi:10.1001/jamanetworkopen.2023.45050.

¹⁷European Parliament and Council of the European Union, Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, 12 July 2024.

¹⁸Cohen, I.G., and Slottje, A., "Artificial Intelligence and the Law of Informed Consent," in Research Handbook on Health, AI and the Law (Cheltenham: Edward Elgar Publishing, 2024), pp. 167-182.

¹⁹Chan, B., "Black-box Assisted Medical Decisions: AI Power vs. Ethical Physician Care," Medicine, Health Care and Philosophy 26, no. 3 (2023): 285-292. doi:10.1007/s11019-023-10153-z.

²⁰Cohen, I.G., and Slottje, A., "Artificial Intelligence and the Law of Informed Consent," in Research Handbook on Health, AI and the Law (Cheltenham: Edward Elgar Publishing, 2024), pp. 167-182.

²¹Amann, J., et al., "Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective," BMC Medical Informatics and Decision Making 20 (2020): 310. doi:10.1186/s12911-020-01332-6.

²²Advarra, "Understanding the Impact of the New EU Artificial Intelligence Act on Clinical Research," Advarra Blog, February 2025. Available at: <https://www.advarra.com/blog/understanding-the-impact-of-the-new-eu-artificial-intelligence-act-on-clinical-research/>.

²³Stibbe, "Legal Considerations for Artificial Intelligence in the Life Sciences Sector," Stibbe Publications and Insights, 2024. Available at: <https://www.stibbe.com>.

²⁴European Medicines Agency and U.S. Food and Drug Administration, "Ten Principles for Good Artificial Intelligence Practice in the Medicines Lifecycle," Joint EMA-FDA Statement, January 2026.

²⁵FDLI, "Regulating the Use of AI in Drug Development: Legal Challenges and Compliance Strategies," Food and Drug Law Institute, July 2025. Available at: <https://www.fdpi.org/2025/07/regulating-the-use-of-ai-in-drug-development/>.

decision-making, patients' rights to human oversight, and any data uses beyond the immediate trial scope.²⁶[5][13]

C. Lifecycle Monitoring and Governance

AI model performance must be continuously monitored throughout the trial and during post-authorisation surveillance. Model drift, changes in performance over time due to shifts in patient populations or clinical practice, represents a significant risk in long-running trials. Sponsors should establish documented procedures for detecting and responding to performance degradation, including pre-specified criteria for model retraining or suspension.²⁸[2] Human-in-the-loop mechanisms, whereby clinical experts retain ultimate decision-making authority over AI-generated recommendations, are a non-negotiable component of lifecycle governance.²⁹[10]

D. Regulatory Engagement and Documentation

Sponsors should engage proactively with regulatory authorities through FDA pre-submission meetings, EMA scientific advice, or qualification opinion procedures, to align on the credibility of AI models used in pivotal trials before submission. Comprehensive technical documentation, including model cards, data governance protocols, and change management logs, should be maintained throughout the trial lifecycle and be available for regulatory inspection.³⁰[9][8]

VI. CONCLUSION

AI models hold genuine transformative potential for clinical trial design and execution. However, this potential can only be responsibly realised if the regulatory and ethical risks associated with AI deployment are rigorously and proactively managed. This paper has identified algorithmic bias and the erosion of meaningful informed consent as the two most pressing ethical risks, and has shown how the evolving regulatory frameworks of the FDA, EMA, and EU AI Act are progressively addressing, though not yet fully resolving, these challenges.

The multi-layered risk management framework proposed here, spanning pre-deployment assessment, ethics review and consent redesign, lifecycle monitoring, and regulatory engagement, offers a practical architecture for sponsors, investigators, and regulators seeking to embed responsible AI governance into clinical research practice. As the FDA-EMA joint principles of January 2026 affirm, international regulatory convergence on AI in drug development is both necessary and underway.³²[4] The responsibility now lies with the research community to translate these principles into operational

governance that places patient safety and scientific integrity at the centre of the AI clinical trial enterprise.

REFERENCES

- [1] U.S. Food and Drug Administration, "Using Artificial Intelligence and Machine Learning in the Development of Drug and Biological Products," Discussion Paper, May 2023, revised February 2025.
- [2] European Medicines Agency, "Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle," EMA/CHMP/CVMP, September 2024.
- [3] European Parliament and Council of the European Union, Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, 12 July 2024.
- [4] European Medicines Agency and U.S. Food and Drug Administration, "Ten Principles for Good Artificial Intelligence Practice in the Medicines Lifecycle," Joint EMA-FDA Statement, January 2026.
- [5] Youssef, A., et al., "Ethical Considerations in the Design and Conduct of Clinical Trials of Artificial Intelligence," JAMA Network Open 7, no. 9 (2024): e2432643.
- [6] Cross, J.L., Choma, M.A., Onofrey, J.A., "Bias in Medical AI: Implications for Clinical Decision-Making," PLOS Digital Health 3, no. 11 (2024): e0000651.
- [7] Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S., "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," Science 366, no. 6464 (2019): 447-453.
- [8] Busch, F., Kather, J.N., Johnner, C., "Navigating the EU AI Act: Implications for Regulated Digital Medical Products," npj Digital Medicine 7 (2024): 210.
- [9] Advarra, "Understanding the Impact of the New EU Artificial Intelligence Act on Clinical Research," Advarra Blog, February 2025.
- [10] Goktas, P., et al., "Shaping the Future of Healthcare: Ethical Clinical Challenges and Pathways to Trustworthy AI," PMC11900311, February 2025.
- [11] Chan, B., "Black-box Assisted Medical Decisions: AI Power vs. Ethical Physician Care," Medicine, Health Care and Philosophy 26, no. 3 (2023): 285-292.
- [12] Amann, J., et al., "Explainability for Artificial Intelligence in Healthcare: A Multidisciplinary Perspective," BMC Medical Informatics and Decision Making 20 (2020): 310.
- [13] Cohen, I.G., Slotte, A., "Artificial Intelligence and the Law of Informed Consent," in Research Handbook on Health, AI and the Law. Edward Elgar Publishing, 2024, pp. 167-182.
- [14] Chin, M.H., et al., "Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Health Care," JAMA Network Open 6 (2023): e2345050.
- [15] Stibbe, "Legal Considerations for Artificial Intelligence in the Life Sciences Sector," Stibbe Publications and Insights, 2024.
- [16] FDLI, "Regulating the Use of AI in Drug Development: Legal Challenges and Compliance Strategies," Food and Drug Law Institute, July 2025.

²⁶Youssef, A., et al., "Ethical Considerations in the Design and Conduct of Clinical Trials of Artificial Intelligence," JAMA Network Open 7, no. 9 (2024): e2432643. doi:10.1001/jamanetworkopen.2024.32643.

²⁷Cohen, I.G., and Slotte, A., "Artificial Intelligence and the Law of Informed Consent," in Research Handbook on Health, AI and the Law (Cheltenham: Edward Elgar Publishing, 2024), pp. 167-182.

²⁸European Medicines Agency, "Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle," EMA/CHMP/CVMP, adopted September 2024. Amsterdam: EMA.

²⁹Goktas, P., et al., "Shaping the Future of Healthcare: Ethical Clinical Challenges and Pathways to Trustworthy AI," Frontiers/PMC, PMC11900311, February 2025.

³⁰Advarra, "Understanding the Impact of the New EU Artificial Intelligence Act on Clinical Research," Advarra Blog, February 2025. Available at: <https://www.advarra.com/blog/understanding-the-impact-of-the-new-eu-artificial-intelligence-act-on-clinical-research/>.

³¹Busch, F., Kather, J.N., and Johnner, C., "Navigating the EU AI Act: Implications for Regulated Digital Medical Products," npj Digital Medicine 7 (2024): 210. doi:10.1038/s41746-024-01213-6.

³²European Medicines Agency and U.S. Food and Drug Administration, "Ten Principles for Good Artificial Intelligence Practice in the Medicines Lifecycle," Joint EMA-FDA Statement, January 2026.