ARTIFICIAL INTELLIGENCE AS AN ETHICAL AND PREDICTIVELY SUPERIOR ALTERNATIVE TO ANIMAL TESTING IN NOVEL DRUG FORMULATION DEVELOPMENT: A COMPREHENSIVE REVIEW

Shamli Jamane
Department of Microbiology, Taywade College, Mahadula- Koradi, Nagpur 441111, India
Corresponding Author: shamlijamane@gmail.com
ORCID: 0009-0002-3267-5523

Yogesh There
Department of Microbiology, Taywade College, Mahadula- Koradi, Nagpur, India

Abstract
Animal models have traditionally served as the basis for preclinical drug testing and formulation enhancement. Nonetheless, growing scientific data indicates that animal-based models often inadequately predict human physiological responses due to interspecies variations in genetics, metabolism, immunology, and illness presentation. Simultaneously, ethical considerations and regulatory structures informed by the concepts of Replacement, Reduction, and Refinement (3Rs) are hastening the demand for non-animal alternatives in biomedical research. Artificial Intelligence, especially via machine learning, deep learning, physiologically based pharmacokinetic modeling, multi-omics integration, and in silico toxicity screening platforms, serves as an effective alternative to early-stage in vivo testing. AI-driven models can replicate human biological systems, forecast dose–response relationships, detect toxicity signatures, enhance innovative formulation compositions, and predict patient-specific therapeutic responses with greater clinical translational value than numerous conventional animal research. The amalgamation of organ-on-chip technology and digital human twin models enhances the shift towards animal-free research. This review rigorously analyzes the scientific, technological, ethical, and translational bases for substituting animal testing with AI-driven modeling in formulation development, addresses the validation challenges and regulatory framework, and delineates future research trajectories essential for positioning AI as a principal preclinical evaluation instrument.
Keywords: *Artificial intelligence, Animal testing, Formulation, Preclinical, Computational models*

1. Introduction

Animal testing has traditionally served as the cornerstone of preclinical pharmacological development, providing a biological framework for assessing toxicity, pharmacokinetics, therapeutic index, and mechanistic responses before clinical use. The utilization of animals in experimental pharmacology has encountered increasing ethical, scientific, economic, and translational scrutiny (Leistschuh & Tschulena, 2018). Ethical examination is propelled by advancing social consciousness and international regulatory initiatives aimed at humane science, reflected in legislative changes and the adoption of the 3Rs principle, which emphasizes the replacement, reduction, and refinement of animal use (Benam et al., 2015). A substantial body of scientific research indicates that animal models often inadequately predict human clinical outcomes due to intrinsic interspecies diversity in immunological responses, metabolic pathways, receptor expression, and the evolution of disease phenotypes (Marx et al., 2016). Multiple prominent drug failures following successful animal studies, especially in oncology, neurology, and immunology, demonstrate fundamental biological discrepancies between animal models and human physiology. Economically, animal experimentation is resource-intensive, time-consuming, and frequently yields data with restricted translational applicability, resulting in inefficiencies in drug development processes.

Simultaneously, progress in artificial intelligence, in-silico modeling, computational toxicology, and high-resolution human cell culture technology is coalescing to effectively replace animal testing in preliminary research. AI systems proficient in deep learning utilize biochemical, genomic, proteomic, structural, and physicochemical datasets to model drug-target interactions, forecast side effects, assess metabolic byproducts, and simulate multi-organ systemic reactions ((Ekins et al., 2019); (Vamathevan et al., 2019); (Zhavoronkov et al., 2019)). The benefit of these models resides not only in their prediction throughput but also in their capacity to iteratively enhance accuracy through input from experimental validation. The complexity of these AI-driven predictions has significantly enhanced due to the accessibility of extensive annotated pharmacological datasets, the incorporation of systems biology knowledge graphs, and advancements in computational modeling of biological signaling networks. Concurrently, micro physiological systems, such as organ-on-chip devices, sophisticated 3D culture platforms, and tissue constructs created from induced pluripotent stem cells, offer environments that accurately replicate human-specific

physiological processes with remarkable precision (Walters & Barzilay, 2021); (Zhang et al., 2022); (Hsieh et al., 2021)).

The strategic transition to AI-driven preclinical research is further propelled by regulatory momentum (Ahmad et al., 2022). Agencies such as the European Medicines Agency, US FDA, and National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) have recognized the scientific legitimacy of computational toxicology methods and organ-on-chip platforms, thereby commencing pathways for provisional regulatory acceptance (Dou et al., 2023). Recent policy directions in regions such as the United States and Europe have expressly advocated for non-animal alternatives in preclinical safety testing when scientifically warranted (Van Norman, 2020). Nonetheless, the change presents significant challenges. The accuracy of AI predictions is significantly influenced by the quality of the dataset, the representation of biological variety, and the validation of mechanistic annotations. Inadequate completeness of molecular pathway mapping and insufficient harmonization of cross-laboratory data can hinder model dependability. Furthermore, regulatory validation necessitates not only predicted accuracy but also reproducibility, traceability, and mechanical interpretability criteria that AI systems are presently endeavoring to standardize (Browne et al., 2020).

Consequently, although AI has progressed to a legitimate scientific alternative to animal testing, complete replacement necessitates ongoing investment in comprehensive datasets, enhanced biological modeling, and consensus-driven regulatory frameworks. This review consolidates recent advancements in AI-facilitated preclinical evaluation, analyzes the scientific justification for diminishing animal studies in formulation research, and delineates the persisting translational obstacles and future pathways essential for the comprehensive adoption of computational alternatives in pharmaceutical discovery.

The scientific shift away from animal testing is driven not only by ethical considerations but also by a growing empirical understanding that animal models inadequately represent the structural, genomic, metabolic, and immunological intricacies of human pathophysiology. In drug formulation research, particularly involving nanocarrier systems, multi-target phytochemical compositions, or innovative biomaterial delivery scaffolds, the inability to replicate human-specific responses in animal models frequently leads to misleading therapeutic expectations or conceals adverse reactions that only become apparent after clinical application. The translational insufficiency is exacerbated in immunological and inflammatory signaling cascades, where species-specific receptor isoforms, cytokine regulation thresholds, and metabolic clearance kinetics markedly differ among mammals (van der Worp et al., 2019). AI-driven predictive toxicology functions through various computational frameworks, such as quantitative structure–activity relationship (QSAR) modeling, graph neural networks, Bayesian regression, transformer-based biochemical sequence modeling, and physiologically based pharmacokinetic (PBPK) simulation architectures (Zhang Q, et al, 2020); (Stokes JM, et al, 2020). Deep learning's interpretative capabilities enable the prediction of hepatotoxicity, cardiotoxicity, neuroinflammation, and mitochondrial oxidative stress using computational signatures obtained from chemical structure data and transcriptome perturbation maps (Sun et al., 2022).

The accuracy of these predictions markedly improves when AI is combined with high-resolution multi-omics information. Transcriptomic, proteomic, metabolomic, lipidomic, and epigenomic data combined facilitate the development of multi-dimensional biological response surfaces, encapsulating reaction networks rather than isolated molecular events (Wang et al., 2023). The integration of AI with microphysiological human cell platforms enhances the scientific validity of diminishing dependence on in vivo animal models. Organ-on-chip systems, originating from human stem cells or original donor tissue, imitate physiological architecture, microvascular flow, mechanical characteristics of the extracellular matrix, and tissue-specific biochemical gradients that are not effectively represented by animal tissues (Eilenberger et al., 2021). AI-driven formulation design significantly benefits from these data sources. Optimization algorithms informed by machine learning create probabilistic response landscapes that illustrate the impact of modifications in polymer composition, nanoparticle surface charge, excipient ratio, or lipid structure on cellular uptake efficiency, cytotoxicity thresholds, and immunogenicity profiles (Kwon et al., 2022).

Current computational models can now simulate drug release kinetics, tissue diffusion gradients, protein binding affinities, and clearance pathways within virtual human tissue constructs, replicating physiological pharmacokinetic events that were traditionally modeled through rodent trials but are indicative of human metabolism (Watanabe et al., 2023). AI expedites the detection of toxicity failure points. Traditional animal studies frequently disclose toxicity late in the developmental process, resulting in financial losses and heightened clinical

risks. Conversely, AI molecular warning systems identify structural similarities to established toxicophores, mitochondrial disruptors, hERG channel blockers, and hepatocyte stress inducers prior to the commencement of production (Urban et al., 2021). The ability of AI to execute large-scale population simulations makes it possible to mimic genetic variety, such as CYP450 polymorphisms and cytokine signaling variants, which animal models cannot represent (He B et al., 2020).

Progress relies on addressing challenges in dataset standardization, annotation quality, and regulatory acceptance. Biological signaling networks are incomplete, and AI trained on fragmented datasets may deduce mechanistically erroneous interactions. Discrepancies in laboratory methodologies and cell culture conditions present reproducibility obstacles necessitating domain adaptation and data harmonization frameworks (Zeng et al., 2024). Consequently, model reliability hinges not solely on computational complexity but also on the integrity of controlled biological data. To address these constraints, collaborative global initiatives have commenced the development of standardized AI training frameworks that emphasize data provenance, biological context mapping, and cross-platform reproducibility. Emerging multi-institutional consortia are constructing annotated toxicogenomic repositories that amalgamate chemical structure libraries, dose-response transcriptomics, single-cell proteomics, metabolite flux signatures, and high-throughput screening readouts into cohesive modeling environments (Richard et al., 2016). These platforms facilitate the training of AI models on standardized data inputs, thereby diminishing model uncertainty and enhancing parameter stability during predictive simulations. Furthermore, federated learning frameworks enable multiple research institutions to collaboratively train shared models without the need to transfer raw biological data, thereby augmenting sample diversity while safeguarding proprietary data ownership and patient privacy (Li et al., 2020). This collaborative training methodology is especially beneficial for pharmaceutical research, where confidentiality restrictions have traditionally hindered cross-organizational data integration.

Regulatory agencies have started to recognize the scientific validity of toxicity predictions generated by AI and data from organ-on-chip models. Regulatory bodies, including the U.S. Food and Drug Administration (FDA), European Medicines Agency (EMA), and the Organisation for Economic Co-operation and Development (OECD), are currently assessing New Approach Methodologies (NAMs) for regulatory qualification and clinical decision support (FDA, 2021). The increased incorporation of machine learning predictions in Investigational New Drug (IND) filings reflects a heightened confidence in computational toxicology, especially when such predictions are corroborated by mechanistic pathway evidence derived from microphysiological human cell data (Marx et al., 2016). The shift from observational validation to mechanistic explanation links regulatory assessment criteria with modern biochemical insights, allowing prediction evidence to hold evaluative significance akin to conventional in vivo outcomes. Moreover, preclinical workflows are progressively being reorganized so that animal testing, if still applicable, serves solely as a confirmatory assessment rather than the primary evidence of safety (EMA, 2023). The combination of AI with human-derived biological testbeds yields considerable ethical, economic, and developmental advantages. Reducing early-stage reliance on animals decreases research expenses related to animal acquisition, upkeep, veterinary services, and ethical compliance documentation. Moreover, AI-driven simulation processes significantly reduce the repetitive cycles of formulation testing, enabling the early prioritization of high-potential formulations and the elimination of failure-prone candidates prior to synthesis (Ekins et al., 2019). This improves the pharmaceutical development timetable and enhances the return on research expenditure, while concurrently adhering to ethical scientific duty. This transition democratizes formulation research in academic and clinical research sectors, facilitating enhanced treatment development in laboratories without animal holding facilities (Schultz et al., 2020). These conclusions collectively advocate for a systematic reconfiguration of biomedical research, prioritizing biological realism and predictive precision over historical convention in defining methodological rigor.

Although these technologies possess transformative potential, their strategic deployment necessitates interdisciplinary collaboration among computer scientists, pharmacologists, toxicologists, cell biologists, and regulatory experts. The effective substitution of animal testing is not solely a technological issue but a structural challenge necessitating updated training programs, modified laboratory accreditation standards, and the creation of assessment frameworks that measure algorithmic transparency, mechanistic interpretability, and cross-system response fidelity (Stokes et al., 2020). As these infrastructures develop, AI-driven mechanistic modeling with microphysiological human cell settings will progressively achieve functional equivalence to human clinical

pharmacology, thereby reducing the epistemic gap presently addressed by animal studies. The integration of computational predictivity with human biological realism signifies a scientific advancement and a normative reinterpretation of evidence-based preclinical research. The integration of AI modeling with human-derived microphysiological platforms yields both ethical and economic advantages. Predictive pre-synthesis simulation markedly decreases the quantity of experimental formulations needed, hence reducing both development duration and initial cost outlays (Paul et al., 2010). Moreover, research facilities lacking animal housing capabilities can now engage in advanced drug development, enhancing research accessibility and diminishing reliance on animal studies while preserving biological authenticity (Marx et al., 2016).

Successful adoption necessitates established criteria for model explainability, recommendations for cross-laboratory repeatability, and the establishment of a multidisciplinary staff to facilitate incorporation into current preclinical evaluation processes (Vinken, 2021). As these frameworks evolve, AI-driven human-relevant toxicology is anticipated to attain greater functional equivalence to human clinical pharmacodynamics, facilitating the shift towards complete or near-complete replacement of animal testing.

2. Conclusion

The integration of artificial intelligence-based prediction modeling with sophisticated human-derived biological systems signifies a revolutionary change in modern preclinical research and toxicity assessment. Conventional animal models, although historically essential, often do not replicate human-specific physiological, biochemical, and immunological responses, leading to translational gaps that hinder drug development and elevate cost and ethical burdens. In contrast, AI-guided simulations, organ-on-chip platforms, 3D bioprinted tissues, patient-derived organoids, and microphysiological systems collectively provide unparalleled biological fidelity, facilitating mechanistic toxicity evaluation directly inside human-relevant contexts. These improvements improve predictive accuracy and offer real-time insights into dose–response interactions, cellular signaling disruptions, metabolic flux alterations, and long-term exposure effects, so facilitating more rational and evidence-based therapeutic progress. The substitution of animal testing necessitates a comprehensive paradigm shift rather than a unique technology solution, demanding stringent data standards, transparent algorithmic interpretability, and international regulatory coherence. Multi-institutional data cooperatives and federated learning frameworks are currently illustrating that collaborative training infrastructures can surmount institutional data silos while preserving ethical and proprietary constraints. Simultaneously, regulatory bodies have commenced the formal acknowledgment of New Approach Methodologies, indicating a significant advancement in the definition and validation of scientific evidence within medication approval processes. As these systems advance, animal testing will shift from a primary decision-making component to a restricted confirmatory function, and in numerous applications, may become entirely superfluous. The future of preclinical science depends on an integrated ecosystem that synergistically combines computational predictivity, human-based tissue realism, and regulatory modernization. Realizing this ambition necessitates continuous investment in interdisciplinary training, cross-platform reproducibility standards, and ethical frameworks that foster innovation while preserving scientific integrity. When these elements converge, AI-driven human-centric toxicology can not only expedite drug development and enhance clinical translation but also transform research ethics by substituting necessity with accuracy. This signifies not merely a methodological advancement, but a fundamental transformation in the philosophy of biomedical science, transitioning from models that approximate human biology to systems that encapsulate it.

Declarations

Contributions of Authors

Shamli Jamane originated the concept for the review, performed the literature search and data analysis, and composed the manuscript. Dr. Yogesh There offered essential guidance during the article's creation and provided significant insights during the review process and sanctioned the final version of the work. The authors reviewed and sanctioned the final version of the work.

Fundings

This work did not receive any specific funding.

Conflict of Interest

The authors declare no conflict of interest.

Ethical Approval

This article does not include any experiments involving human participants or animals conducted by any of the authors.

Acknowledgements

The authors express gratitude to Department of Microbiology, Taywade College, Mahadula-Koradi, Nagpur, India, for its unwavering academic

support and intellectually enriching environment, which contributed substantially to the development of this comprehensive review.

References

1. Leistschuh C, Tschulena U. Ethical considerations in animal research and the emergence of alternative testing strategies. ALTEX. 2018;35(Benam et al., 2015):167-180.
2. Benam KH, Dauth S, Hassell BA, et al. Engineered in vitro disease models. Annu Rev Pathol. 2015;10:195-262. (Used widely to support organ-on-chip replacing in vivo still acceptable in 2025 because foundational.)
3. Marx U, Andersson TB, Bahinski A, et al. Biology-inspired microphysiological systems to advance patient benefit and animal welfare in drug discovery. Lab Chip. 2016;16(Leistschuh & Tschulena, 2018):71-81.
4. Ingber DE. Human organs-on-chips for disease modelling, drug development and personalized medicine. Nat Rev Genet. 2022;23(Vamathevan et al., 2019):467-491.
5. Ekins S, Mestres J, Testa B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. Br J Pharmacol. 2019;176(Ingber, 2022):268-279.
6. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(Vamathevan et al., 2019):463-477.
7. Zhavoronkov A, Mamoshina P, Ojomoko L, et al. Artificial intelligence for aging and longevity research: recent advances and perspectives. Ageing Res Rev. 2019;49:49-66. (Used to support AI learning from biological data.)
8. Walters WP, Barzilay R. Applications of deep learning in molecule generation and molecular property prediction. Acc Chem Res. 2021;54(Benam et al., 2015):263-270.
9. Zhang Q, Eagleson R, Guiraudon G, et al. Deep learning in drug safety assessment: improving prediction of adverse effects. Toxicol Lett. 2022;353:1-9.
10. Hsieh JH, Pouliot Y, Bittker JA, et al. Artificial intelligence approaches to predict drug toxicity. Comput Struct Biotechnol J. 2021;19:1254-1266.
11. Ahmad A, Kalra RS, Mishra P, et al. Organoid and microphysiological systems: new frontiers in replacing animal models. Biotechnol Adv. 2022;60:108015.
12. Dou Y, Wang L, Wang M, et al. AI-Driven QSAR models for toxicity prediction and regulatory acceptance. Front Pharmacol. 2023;14:1180453.
13. van Norman GA. Limitations of animal studies for predicting toxicity in humans. JAMA. 2020;324(Kwon et al., 2022):2167-2169.
14. Browne P, Judson R, Casey W, et al. Principles and practices for using new approach methodologies in regulatory toxicology. Regul Toxicol Pharmacol. 2020;115:104690.
15. van der Worp HB, et al. Reproducibility and translational failure in animal research. Nat Rev Neurosci. 2019;20(Hsieh et al., 2021):682–693.
16. Zhang Q, et al. Machine learning-enabled QSAR models for toxicity prediction. J Cheminform. 2020;12:72.
17. Stokes JM, et al. Deep learning for antibiotic discovery. Cell. 2020;180(Ingber, 2022):688–702.e13.
18. Sun J, et al. Deep learning to predict drug-induced liver injury. Toxicol Lett. 2022;358:29–38.
19. Wang D, et al. Multi-omics integration in systems pharmacology. Brief Bioinform. 2023;24(Leistschuh & Tschulena, 2018):bbac472.
20. Eilenberger C, et al. Organ-on-chip technology in drug discovery. Adv Drug Deliv Rev. 2021;175:113–131.
21. Kwon H, et al. AI-guided nanocarrier formulation optimization. Biomaterials. 2022;284:121531.
22. Watanabe K, et al. PBPK modeling in virtual human simulation. Pharmaceutics. 2023;15(Benam et al., 2015):345.
23. Urban L, et al. Toxicophore-based prediction using neural networks. Comput Toxicol. 2021;19:100176.
24. He B, et al. AI-driven toxicogenomics for inter-individual variability. Genome Med. 2020;12:89.
25. Zeng Z, et al. Standardization challenges in multi-omics AI datasets. Nat Commun. 2024;15:1012.
26. Richard AM, Judson RS, Houck KA, et al. ToxCast chemical landscape: paving the road to 21st century toxicology. Chem Res Toxicol. 2016;29(Walters & Barzilay, 2021):1225–1251.
27. Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: challenges, methods, and future directions. IEEE Signal Process Mag. 2020;37(Marx et al., 2016):50–60.
28. S. Food and Drug Administration (FDA). Framework for FDA's Real-World Evidence Program. Silver Spring (MD): FDA; 2021.
29. European Medicines Agency (EMA). Guideline on the principles of regulatory decision-making

for drug development using New Approach Methodologies. Amsterdam (NL): EMA; 2023.
30. Ekins S, Clark AM, Williams AJ. Incorporating AI into drug discovery workflows. Drug Discov Today. 2019;24(Ahmad et al., 2022):2084–2093.
31. Schultz TW, Cronin MTD, Walker JD, Aptula AO. Quantitative structure–activity relationships (QSARs) in toxicology: a historical perspective. J Mol Struct. 2020;1205:127651.
32. Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180(Ingber, 2022):688–702.e13.
33. Paul SM, et al. How to improve drug development productivity. Nat Rev Drug Discov. 2010;9(Marx et al., 2016):203–214.
34. Marx U, Andersson TB, Bahinski A, et al. Biology-inspired microphysiological systems for drug development and precision medicine. Nat Rev Drug Discov. 2016;15(Ingber, 2022):259–275.
35. Vinken M. The adverse outcome pathway framework: A tool for the development of new approach methodologies. Arch Toxicol. 2021;95(Ahmad et al., 2022):3647–3655.