Int. j. adv. multidisc. res. stud. 2026; 6(2):1170-1177

Received: 17-02-2026
Accepted: 27-03-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Explainability, Regulatory Compliance, and Ethical Dimensions of Artificial Intelligence in Pharmaceutical Research and Development

[1] Nithya Shumedha, [2] Nawaz Mahammed

[1, 2] Department of Pharmaceutics, Raghavendra Institute of Pharmaceutical Education and Research, KR Palli Cross, Anantapur, Chiyyedu, Andhra Pradesh, India

DOI: https://doi.org/10.62225/2583049X.2026.6.2.6081

Corresponding Author: **Nawaz Mahammed**

Abstract

In recent years, Artificial Intelligence (AI) and Machine Learning (ML) have become transformative tools across pharmaceutical research and development spanning from target discovery, drug design, formulation, to clinical decision support and post-marketing surveillance. However, the widespread adoption of AI in regulated pharmaceutical settings is constrained by critical issues of explainability, regulatory compliance, and ethical governance. Without interpretability or transparency, AI "black boxes" risk undermining trust, raising safety concerns, and impeding regulatory acceptance. Regulators such as the U.S. Food and Drug Administration (FDA) are evolving guidance frameworks (e.g., for AI/ML in drug development and Medical Device Software) that emphasize credibility, life-cycle management, and documentation. Meanwhile, ethical dimensions such as bias, accountability, data privacy, and informed consent require careful integration into AI deployment strategies. In this review, we synthesize the state of explainable AI (XAI) methods tailored to pharmaceutical applications, map current regulatory landscapes (with emphasis on FDA, EMA, and emerging Indian regulation), and discuss ethical considerations unique to drug development. We propose a unified framework to guide researchers and industry in deploying AI systems that are interpretable, compliant, and ethical thereby enabling safer, trustworthy adoption of AI in pharmaceutical R&D.

Keywords: Artificial Intelligence, Explainability, Regulatory Compliance, Drug Development

1. Introduction

1.1 Background

Artificial Intelligence (AI) and Machine Learning (ML) are transforming pharmaceuticals by driving data-based decisions across drug discovery, development, and post-marketing surveillance [1]. AI identifies novel targets, predicts properties, optimizes formulations, and supports clinical trials and pharmacovigilance through real-world data analysis. These technologies accelerate development, reduce costs, and enhance precision medicine. However, their integration raises challenges in data reliability, transparency, and accountability within regulated environments. As AI increasingly influences regulatory and clinical outcomes, ensuring explainability, ethical integrity, and regulatory compliance is crucial to build trust and guarantee the safe, effective, and responsible use of AI in pharmaceutical research and innovation [2, 3]

1.2 The Problem: "Black-box" nature of AI systems and associated risks

AI's "black-box" nature poses risks in pharmaceutical R&D, as opaque models hinder transparency, reproducibility, and regulatory trust. Without explainability, biased or unverifiable predictions can endanger patient safety and equity. Integrating explainable AI, ethical oversight, and rigorous validation is essential to ensure compliant, transparent, and scientifically credible AI systems [4, 5].

1.3 Why Explainability, Regulation, and Ethics Matter

Explainability, regulation, and ethics underpin responsible AI in pharmaceuticals. Explainability builds trust, regulation ensures safety and compliance, and ethics safeguards fairness and patient rights. Their integration prevents bias, data misuse, and unsafe decisions, fostering trustworthy, compliant, and socially responsible AI-driven pharmaceutical innovation [6].

1.4 Aim of the Review

This review synthesizes insights on explainability, ethics, and regulatory compliance in pharmaceutical AI, evaluating XAI methods, global frameworks, and proposing an integrated model for transparent, ethical, and regulatory-ready AI adoption in drug development.

2. Role of AI in Pharmaceutical R&D

2.1 AI in Drug Discovery, Preclinical, and Clinical Research

AI is revolutionizing every phase of pharmaceutical research, from target discovery to clinical validation. In drug discovery, AI algorithms analyze large-scale omics and chemical data to identify novel therapeutic targets and predict molecular interactions with remarkable efficiency [7]. Techniques such as deep learning, graph neural networks, and generative models assist in virtual screening and de novo molecule design, reducing time and cost [8]. During preclinical research, AI supports bioactivity prediction, ADMET profiling, and in silico toxicology modelling, minimizing animal testing. In clinical development, machine learning aids in patient recruitment, adaptive trial design, and real-time data analysis to enhance trial success rates [9]. AI-driven clinical decision systems also identify biomarkers, optimize dosing, and predict adverse responses. Collectively, these applications accelerate drug development pipelines, reduce attrition, and promote personalized therapeutics marking a shift from hypothesis-driven to data-driven pharmaceutical innovation [10].

2.2 AI-Driven Formulation Development, Toxicity Prediction, and Pharmacovigilance

AI plays an increasingly vital role in formulation development by modelling complex relationships between formulation variables and performance outcomes [11]. Machine learning predicts dissolution, stability, and bioavailability, guiding the optimization of dosage forms and nanocarriers [12]. In toxicity prediction, AI-based QSAR and deep learning models forecast potential adverse effects before clinical testing, saving significant resources and improving safety. These models integrate chemical structure, biological interaction, and omics data to predict hepatotoxicity, cardiotoxicity, or genotoxicity. Moreover, AI in pharmacovigilance enhances post-marketing surveillance by detecting adverse drug reactions (ADRs) from electronic health records, social media, and real-world databases through natural language processing. This automated detection facilitates rapid regulatory response and risk minimization. Overall, AI unifies predictive modelling, mechanistic understanding, and real-world monitoring offering a robust, data-driven foundation for safer, more efficient pharmaceutical development and lifecycle management [13].

2.3 Opportunities vs. Challenges: Accuracy vs. Interpretability; Innovation vs. Regulation

While AI offers unprecedented accuracy and efficiency, it also introduces critical challenges regarding interpretability and regulatory acceptance. Highly complex deep learning models can outperform traditional statistical approaches but often function as opaque "black boxes," making it difficult for scientists and regulators to understand their reasoning [14]. This lack of explainability hinders trust and validation, particularly when AI decisions influence clinical outcomes or regulatory submissions. Similarly, the rapid innovation in AI-driven methods often outpaces the development of regulatory frameworks designed to ensure safety, fairness, and transparency [15]. Achieving equilibrium between innovation and compliance requires standardized validation protocols, lifecycle monitoring, and explainable algorithms. Additionally, ethical issues such as data bias, privacy, and accountability must be addressed. Thus, while AI presents vast opportunities to transform pharmaceutical R&D, its sustainable implementation depends on balancing technological sophistication with transparency, governance, and ethical responsibility [16].

3. Explainability In Pharmaceutical AI Systems

3.1 Understanding Explainable AI (XAI)

Explainable Artificial Intelligence (XAI) refers to a collection of techniques and methodologies designed to make the inner workings of AI systems transparent, interpretable, and understandable to human users [17]. In the pharmaceutical domain, XAI ensures that AI-generated decisions such as drug target predictions, toxicity assessments, or patient risk stratifications can be rationalized and validated scientifically. The key principles of XAI include transparency (clarity on how inputs influence outputs), interpretability (the ability of humans to comprehend model reasoning), and causality (identifying cause–effect relationships rather than mere correlations) [18]. These principles are essential in regulated environments, where every algorithmic decision must withstand scrutiny by researchers and regulators. XAI bridges the gap between computational intelligence and human reasoning, providing insights into model logic, confidence levels, and reliability [19]. Ultimately, explainable AI fosters trust, accountability, and reproducibility ensuring that advanced machine learning models can be safely integrated into pharmaceutical R&D and clinical decision-making.

3.2 Techniques for Explainability

Explainability techniques help interpret how AI and ML models make predictions, especially in high-stakes pharmaceutical contexts where decisions affect patient safety and regulatory approval [20]. These techniques are categorized broadly into post-hoc (applied after model training) and intrinsic (built into model architecture). This section focuses on post-hoc explainability methods widely used in pharmaceutical research and healthcare analytics [21]. Post-hoc methods do not alter the underlying model; instead, they analyze model outputs to infer reasoning. They are crucial for black-box models like neural networks or ensemble learning algorithms commonly used in drug discovery, toxicity prediction, and clinical data modeling.

3.2.1 Local Interpretable Model-Agnostic Explanations (LIME)

LIME is a post-hoc explainability method that interprets individual predictions of any machine learning model by creating a simple, local surrogate model around a specific data instance [22]. It perturbs input features, observes resulting output changes, and fits an interpretable model—such as linear regression to approximate the original model's local behavior. In pharmaceutical research, LIME helps explain why a molecule is classified as toxic or active, improving understanding of ADMET or QSAR models [23]. Although intuitive and model-agnostic, its explanations are

valid only locally and may vary with sampling, limiting its stability across different runs or datasets.

3.2.2 SHapley Additive exPlanations (SHAP)

SHAP applies cooperative game theory to quantify each feature's contribution to a model's prediction [24]. It computes Shapley values the average marginal effect of including each feature across all possible feature combinations providing both global and local interpretability. Widely used in pharmacoinformatics, SHAP explains molecular property predictions, identifies critical biological determinants of efficacy, and supports regulatory transparency [25]. Its strong theoretical foundation and consistency make it suitable for FDA-aligned explainability workflows. However, SHAP is computationally intensive for high-dimensional pharmaceutical data and assumes feature independence, requiring expert interpretation of plots to avoid misrepresentation of complex feature interactions [26].

3.2.3 Gradient-weighted Class Activation Mapping (Grad-CAM)

Grad-CAM is a visualization technique used in convolutional neural networks (CNNs) to highlight regions of an image that most influence the model's decision [27]. By computing gradients of the output with respect to convolutional feature maps, Grad-CAM generates heatmaps showing where the model focuses during classification [28]. In pharmaceutical applications, it aids in interpreting histopathology or microscopy images, validating drug response models, and ensuring biologically meaningful learning. Grad-CAM provides intuitive, visual explanations valuable for regulatory audits and clinician review. Nonetheless, it is restricted to CNN-based image models and produces qualitative, coarse explanations unsuitable for non-visual data [29, 30].

3.3 Importance in Pharma

Explainability holds special significance in the pharmaceutical sector, where decisions directly impact patient safety and regulatory approval. In toxicity prediction, XAI allows researchers to trace back which chemical substructures or molecular descriptors lead to a model's classification of a compound as toxic, thereby improving confidence and enabling rational drug redesign [31]. In clinical trials, interpretable AI supports patient stratification by revealing which biomarkers or clinical parameters influence therapeutic outcomes, enhancing personalized medicine approaches [32]. Moreover, XAI aids in validating AI models used for dose optimization, adverse event prediction, or pharmacogenomic analyses. Regulatory agencies increasingly expect transparency in algorithmic reasoning before accepting AI-generated data for drug approval processes [31, 32]. By making model outputs scientifically explainable, XAI bridges the trust gap between data scientists, clinicians, and regulators transforming AI from a predictive "black box" into a trustworthy decision-support system capable of improving safety, reproducibility, and ethical accountability in pharma [33, 34].

3.4 Case Studies

Several successful case studies demonstrate how explainable AI has improved transparency and trust in pharmaceutical research [35, 36]. For example, AstraZeneca integrated SHAP-based interpretability within its deep learning pipeline for toxicity prediction, allowing chemists to visualize how specific molecular fragments contributed to hepatotoxic outcomes, leading to safer compound selection [37]. Similarly, Pfizer used interpretable machine learning models in clinical trial recruitment, identifying patient features most predictive of treatment response while ensuring fairness and compliance with regulatory standards [38]. In pharmacovigilance, explainable NLP systems have been applied to spontaneous adverse event reports, clarifying which textual features signaled potential safety concerns [39]. These XAI implementations not only enhanced scientific insight but also facilitated regulatory acceptance by providing auditable explanations of AI decisions. Such case studies illustrate how transparency transforms AI from a mere analytical tool into a credible partner in evidence-based pharmaceutical development and regulatory review.

3.5 Challenges

Despite its promise, XAI faces several challenges in pharmaceutical applications. The most critical is the trade-off between model performance and interpretability: highly accurate deep neural networks often lack transparency, whereas interpretable models like decision trees may underperform on complex biomedical data [40]. Moreover, there is a lack of standardization in how explainability is measured, reported, and validated across studies. Regulatory agencies have yet to define universally accepted metrics for AI interpretability, making cross-comparison difficult [41]. Another challenge lies in contextual misinterpretation human users may overtrust or misread explanations, leading to erroneous conclusions. Additionally, XAI methods are often computationally intensive and require domain expertise to interpret correctly [42]. The absence of unified reporting frameworks and explainability benchmarks hinders reproducibility and regulatory harmonization [43]. Overcoming these challenges will require interdisciplinary collaboration, standard guidelines, and education to balance interpretability, accuracy, and usability in future pharmaceutical AI systems.

4. Regulatory Compliance Landscape for AI in Pharmaceuticals

4.1 Overview of Global Regulatory Agencies

Global regulatory bodies are advancing frameworks for AI governance in pharmaceuticals. The FDA promotes Good Machine Learning Practice (GMLP) and regulates AI/ML-based SaMD for safety and transparency. The EMA's 2023 Reflection Paper stresses traceability and human oversight, while India's CDSCO develops ethical AI policies under the Digital Health Mission collectively promoting standardized, responsible, and transparent AI regulation worldwide.

4.2 Data Integrity and Validation

Data integrity underpins regulatory compliance in pharmaceutical AI. Regulations like 21 CFR Part 11 and GxP ensure secure, traceable, and auditable data use, while the ALCOA+ principles (Attributable, Legible, Contemporaneous, Original, Accurate, etc.) mandate accuracy and reliability. Maintaining data lineage, version control, and audit trails ensures trustworthy, compliant AI systems that uphold patient safety and regulatory confidence.

4.3 Model Validation & Lifecycle Management

Regulatory compliance mandates rigorous AI model

validation to ensure accuracy, reproducibility, and robustness. Agencies require continuous lifecycle management, with revalidation after updates and detailed audit trails for traceability. Ongoing monitoring, version control, and change management under GMLP and GxP frameworks uphold accountability and reliability, enabling sustained regulatory acceptance of pharmaceutical AI systems.

4.4 Regulatory Expectations for Explainability

Regulatory bodies like the FDA and EMA require AI models in drug development to be transparent and auditable, with clear justification of predictions. Explainability and human-in-the-loop oversight ensure scientific accountability and safety. Black-box models lacking interpretability risk rejection, as regulators prioritize transparency, bias mitigation, and verifiable decision-making throughout the pharmaceutical AI lifecycle.

4.5 Emerging International Guidelines

Global bodies like WHO, OECD, and ICH are shaping frameworks for ethical, transparent, and trustworthy AI in pharmaceuticals. WHO promotes human-centered AI, OECD advocates fairness and accountability, and ICH integrates AI into data integrity and clinical guidelines. These initiatives foster global regulatory convergence, ensuring validated, unbiased, and ethically compliant AI systems across healthcare and drug development.

5. Ethical Dimensions of AI in Pharmaceutical Research

5.1 Core Ethical Principles

AI in pharmaceuticals must uphold beneficence, non-maleficence, autonomy, and justice. It should enhance safety and efficacy while preventing harm through unbiased, validated decisions [44]. Respecting data ownership and consent ensures autonomy, while justice promotes equitable access and fairness. Embedding these ethics fosters patient-centered, compliant, and socially responsible AI aligned with healthcare values, not just technical efficiency [45].

5.2 Algorithmic Bias and Fairness

Algorithmic bias in pharmaceutical AI arises from imbalanced or underrepresented datasets, leading to inaccurate or inequitable predictions in drug response or toxicity [46]. Ensuring fairness requires diverse data, bias detection metrics, and ethical oversight. Transparent reporting and cross-population validation promote accuracy, equity, and public trust, ensuring AI-driven pharmaceutical research remains scientifically valid and ethically responsible [47].

5.3 Data Privacy and Confidentiality

Patient data use in pharmaceutical AI demands strict privacy protection under GDPR, HIPAA, and India's DPDPA (2023) [48]. These laws ensure informed consent, data minimization, and secure handling through anonymization and encryption. Embedding privacy by design safeguards patient dignity and trust, making ethical data protection both a regulatory requirement and a moral responsibility in AI research [49].

5.4 Accountability and Human Oversight

Accountability in pharmaceutical AI requires clear responsibility for AI-driven outcomes, with human oversight central to decision-making [50]. The human-in-the-loop approach ensures AI assists rather than replaces experts [51]. Transparent governance, audits, and documentation of model logic and limitations uphold traceability, ethical integrity, and legal defensibility, preventing liability gaps and ensuring safe, accountable AI deployment [52].

5.5 Ethical Use of Patient Data and Clinical Trial Information

Ethical AI in pharmaceuticals requires informed consent, data anonymization, and adherence to guidelines like the Declaration of Helsinki and ICMR GCP [53]. Oversight committees must ensure data is used only for approved purposes. Ethically managed datasets protect participant rights, ensure confidentiality, and strengthen trust, data quality, and model reliability in pharmaceutical research [54].

5.6 Transparency and Public Trust

Transparency in pharmaceutical AI builds public trust through clear communication of model development, validation, and limitations [55]. Openly sharing data sources, biases, and explainability analyses promotes accountability, scientific integrity, and ethical compliance, ensuring AI-driven innovations remain safe, equitable, and socially trustworthy in healthcare [56].

6. Integrating Explainability, Regulation, and Ethics: A Unified Framework

Integrating explainability, regulatory compliance, and ethical governance ensures transparent, responsible, and trustworthy pharmaceutical AI [57]. This section presents a triangular model, stepwise framework, and governance checklist to align innovation with safety, accountability, and public trust **as shown in Table 1 and Fig 1**.

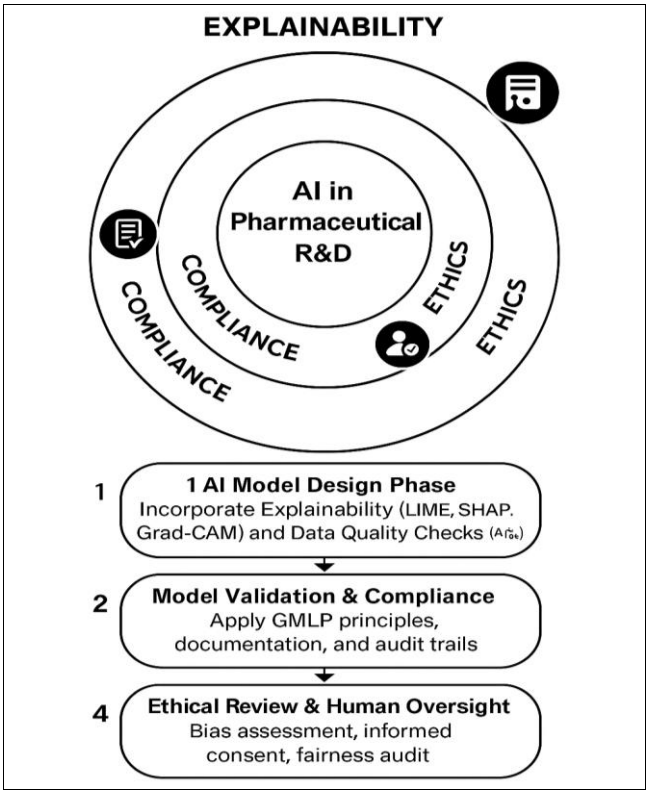


Fig 1: Unified Framework for Explainable, Ethical, and Regulatory-Compliant AI in Pharmaceutical R&D

Table 1: Unified Framework for Responsible AI in Pharmaceuticals

Component	Focus Area	Key Objectives	Expected Outcome
Explainability	Model Transparency	Enhance interpretability and traceability of AI decisions	Builds trust and scientific validation
Regulatory Compliance	Data Integrity & Validation	Ensure adherence to FDA, EMA, CDSCO, GxP, and ALCOA+ standards	Facilitates regulatory acceptance
Ethical Governance	Fairness & Accountability	Prevent bias, protect privacy, and ensure informed consent	Promotes patient safety and equity
Integration Framework	Design–Validation–Deployment Cycle	Harmonize technical, ethical, and legal requirements	Establishes trustworthy AI ecosystem
Governance Practices	Oversight & Monitoring	SOPs, audits, and cross-functional evaluations	Continuous ethical and regulatory assurance

7. Case Studies and Current Industry Practices

AI has become a cornerstone of pharmaceutical research, with global companies adopting explainable, ethical, and compliant systems across discovery to pharmacovigilance, balancing innovation with transparency, governance, and regulatory alignment [58] as shown in Table 2.

Table 2: Real-World Case Studies on AI Implementation in the Pharmaceutical Industry

Company / Institution	AI Application Area	Explainability / XAI Practices	Ethical or Regulatory Integration	Impact / Outcome
AstraZeneca	AI-based predictive toxicology and target discovery using deep learning models.	Used SHAP (Shapley Additive Explanations) to visualize molecular fragments influencing toxicity outcomes.	Followed internal AI governance charter ensuring algorithmic fairness and traceable decision-making.	Improved early-stage compound rejection accuracy by ~30%, reducing animal testing and costs.
Pfizer	AI in clinical trial optimization – patient selection, dosing, and outcome prediction.	Applied interpretable ML models highlighting key biomarkers for patient stratification.	Models validated under GxP guidelines; used human-in-the-loop oversight during trial design.	Reduced patient dropout rates and improved trial efficiency by 15–20%.
Novartis	AI-driven drug discovery via collaboration with Microsoft using explainable deep learning pipelines.	Implemented attention-based neural networks to highlight chemical features influencing activity predictions.	Developed internal "Responsible AI Framework" aligned with EMA transparency recommendations.	Accelerated hit identification and improved decision traceability during molecule screening.
GSK (GlaxoSmithKline)	Predictive modeling for vaccine design and antigen response prediction.	Used model interpretability dashboards to visualize immune response features.	Complied with WHO ethical standards for health data and privacy-by-design principles.	Enhanced predictive accuracy for immune responses; reduced preclinical experimentation timelines.
Sanofi	AI-supported pharmacovigilance analyzing adverse drug reactions from post-marketing reports.	Integrated NLP explainability to highlight key phrases and causal patterns in ADR reports.	Used internal AI ethics review board; maintained audit trails for regulatory audit.	Increased ADR detection sensitivity by 40%, improving post-market safety monitoring.
BenevolentAI (UK)	Knowledge graph–based drug repurposing for diseases like ALS and COVID-19.	Deployed transparent graph reasoning engines for traceable predictions.	Open-source documentation of algorithms ensures reproducibility and external validation.	Identified multiple new drug candidates rapidly; several progressed to clinical trials.
Roche / Genentech	AI for digital pathology and clinical image interpretation.	Employed Grad-CAM and saliency maps to interpret deep learning predictions.	Validated AI algorithms under FDA's SaMD framework with continuous revalidation processes.	Improved diagnostic reproducibility and pathologist decision support in oncology trials.
Google DeepMind & Isomorphic Labs	AI-driven protein structure prediction (AlphaFold) and molecular design.	Publicly released explainable model outputs and confidence metrics for predictions.	Shared data under ethical open-science agreements promoting transparency and collaboration.	Revolutionized structure-based drug design; reduced experimental costs and timelines drastically.
Bayer Pharmaceuticals	AI in manufacturing and process optimization.	Applied interpretable process analytics models for batch quality prediction.	Adhered to GxP data integrity and lifecycle validation under 21 CFR Part 11.	Improved production consistency and reduced deviations through predictive monitoring.
Takeda Pharmaceuticals	AI-based real-world evidence (RWE) analysis from patient registries.	Used explainable NLP models to interpret outcome drivers in text-based medical data.	Conducted ethical impact assessments; ensured patient data anonymization per GDPR.	Improved understanding of treatment outcomes and post-market effectiveness.

8. Conclusion

Artificial Intelligence is redefining pharmaceutical research, yet its responsible integration depends on explainability, regulatory compliance, and ethical governance. This review highlights how transparent, interpretable, and validated AI models foster regulatory trust and patient safety. Establishing unified frameworks such as Good Machine Learning Practice (GMLP) and ethical oversight ensures accountability, fairness, and reliability. Global harmonization among agencies like FDA, EMA, and CDSCO will be vital to standardize AI practices. Ultimately, sustainable pharmaceutical innovation requires balancing technological advancement with transparency, ethics, and compliance transforming AI into a trustworthy catalyst for safe, equitable, and effective healthcare solutions.

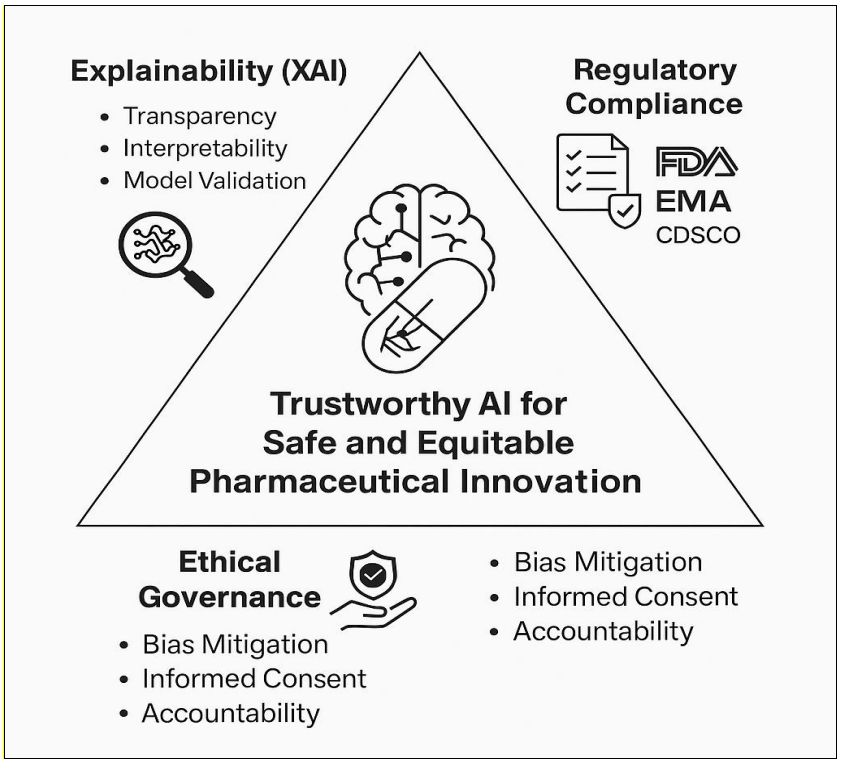
9. Acknowledgments

The authors of extremely thankful to management of Raghavendra Institute of Pharmaceutical Education and Research (RIPER) – Autonomous.

10. Conflicts of Interest

The authors declare no conflicts of interest related to the content of this manuscript. All opinions and conclusions expressed are solely those of the authors and do not reflect the views of their affiliated institutions.

Graphical Abstract:



11. References

1. Mandle N, Rathor S. AI-Enabled Devices in Drug Discovery: Bridging the Gap Between Research and Clinical Application. Journal of Pharmaceutical Research and Integrated Medical Sciences (JPRIMS). 2025; 2(2):25-42.
2. Tiwari PC, Pal R, Chaudhary MJ, Nath R. Artificial intelligence revolutionizing drug development: Exploring opportunities and challenges. Drug Development Research. 2023; 84(8):1652-1663.
3. Chen Z, Liu X, Hogan W, Shenkman E, Bian J. Applications of artificial intelligence in drug development using real-world data. Drug discovery today. 2021; 26(5):1256-1264.
4. Schnitzer S, Tiruwuha MH. Transparency in the Black Box, 2024.
5. Kanagarajah A. AI-driven innovation in healthcare product development: Challenges and ethical implications, 2024.
6. Moorthy UMK, Muthukumaran AMJ, Kaliyaperumal V, Jayakumar S, Vijayaraghavan KA. Explainability and Regulatory Compliance in Healthcare: Bridging the Gap for Ethical XAI Implementation. Explainable Artificial Intelligence in the Healthcare Industry, 2025, 521-561.
7. Bhatt P, Singh S, Kumar V, Nagarajan K, Mishra SK, Dixit PK, *et al*. Artificial intelligence in pharmaceutical industry: Revolutionizing drug development and delivery. Current Artificial Intelligence. 2024; 2(1):E051223224198.
8. Cheng Y, Gong Y, Liu Y, Song B, Zou Q. Molecular design in drug discovery: A comprehensive review of deep generative models. Briefings in bioinformatics. 2021; 22(6):bbab344.
9. Anuyah S, Singh MK, Nyavor H. Advancing clinical trial outcomes using deep learning and predictive modelling: Bridging precision medicine and patient-centered care, 2024. arXiv preprint arXiv:241207050
10. Ozaybi MQB, Madkhali ANM, Alhazmi MAM, Faqihi HMA, Alanazi MM, Siraj WHY, *et al*. The Role of Artificial Intelligence in Drug Discovery and

Development. Egyptian Journal of Chemistry. 2024; 67(13):1541-1547.
11. Ali KA, Mohin S, Mondal P, Goswami S, Ghosh S, Choudhuri S. Influence of artificial intelligence in modern pharmaceutical formulation and drug development. Future Journal of Pharmaceutical Sciences. 2024; 10(1):53.
12. Bannigan P, Aldeghi M, Bao Z, Häse F, Aspuru-Guzik A, Allen C. Machine learning directed drug formulation development. Advanced Drug Delivery Reviews. 2021; 175:113806.
13. Vignesh M. Al in Predictive. AI-Powered Advances in Pharmacology. 2024; 61.
14. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature machine intelligence. 2019; 1(5):206-215.
15. Loveth RG, Covenant L. Regulatory Challenges for AI-Driven Innovations, 2024.
16. Alabi M. Ethical Challenges in AI: Addressing Bias, Privacy, and Accountability, 2025.
17. Minh D, Wang HX, Li YF, Nguyen TN. Explainable artificial intelligence: A comprehensive review. Artificial Intelligence Review. 2022; 55(5):3503-3568.
18. Chinnaraju A. Explainable AI (XAI) for trustworthy and transparent decision-making: A theoretical framework for AI interpretability. World Journal of Advanced Engineering Technology and Sciences. 2025; 14(3):170-207.
19. Ozdemir O, Fatunmbi TO. Explainable AI (XAI) in Healthcare: Bridging the Gap between Accuracy and Interpretability. Journal of Science, Technology and Engineering Research. 2024; 2(1):32-44.
20. Singh S, Yadav D. Explainable AI: Bridging Performance, Transparency, and Accountability in High-Stakes Systems. Next-Gen Machine Learning: Algorithms for Adaptive Intelligence. 84.
21. Bhati D, Amiruzzaman M, Zhao Y, Guercio A, Le T. A Survey of Post-Hoc XAI Methods From a Visualization Perspective: Challenges and Opportunities. IEEE Access, 2025.
22. Sokol K, Hepburn A, Santos-Rodriguez R, Flach P. bLIMEy: Surrogate prediction explanations beyond LIME, 2019. arXiv preprint arXiv:191013016
23. Qadri YA, Shaikh S, Ahmad K, Choi I, Kim SW, Vasilakos AV. Explainable Artificial Intelligence: A Perspective on Drug Discovery. Pharmaceutics. 2025; 17(9):1119.
24. Zhao C, Liu J, Parilina E. The Shapley Value Contribution to Explainable Artificial Intelligence: A Comprehensive Survey. Dynamic Games and Applications, 2025, 1-38.
25. Chen H, Covert IC, Lundberg SM, Lee S-I. Algorithms to estimate Shapley value feature attributions. Nature Machine Intelligence. 2023; 5(6):590-601.
26. Kovvuri V. Explainable Artificial Intelligence across Domains: Refinement of SHAP and Practical Applications, 2024.
27. Chattopadhay A, Sarkar A, Howlader P, Balasubramanian VN, editors. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. 2018 IEEE winter conference on applications of computer vision (WACV), IEEE, 2018.
28. Kumar S, Abdelhamid AA, Tarek Z. Visualizing the unseen: Exploring GRAD-CAM for interpreting convolutional image classifiers. J Full Length Artic. 2023; 4:34-42.
29. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, *et al*. Applications of machine learning in drug discovery and development. Nature reviews Drug discovery. 2019; 18(6):463-477.
30. Raghavan K, Balasubramanian S, Veezhinathan K. Explainable artificial intelligence for medical imaging: Review and experiments with infrared breast images. Computational Intelligence. 2024; 40(3):e12660.
31. Chandak T, Jayashree J, Vijayashree J. Trends and Advancements of AI and XAI in Drug Discovery. Explainable AI (XAI) for sustainable development: Chapman and Hall/CRC, 2024, 233-255.
32. Kumar A, Metta D. AI-driven precision oncology: Predictive biomarker discovery and personalized treatment optimization using genomic data. Int J Adv Res Publ Rev. 2024; 1(3):21-38.
33. Tuan DA. Bridging the gap between black box AI and clinical practice: Advancing explainable AI for trust, ethics, and personalized healthcare diagnostics, 2024.
34. Akande OA. Leveraging explainable AI models to improve predictive accuracy and ethical accountability in healthcare diagnostic decision support systems, 2020.
35. Srinivasu PN, Sandhya N, Jhaveri RH, Raut R. From blackbox to explainable AI in healthcare: Existing tools and case studies. Mobile Information Systems. 2022; 2022(1):8167821.
36. Rosenbacke R, Melhus Å, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. Jmir Ai. 2024; 3:e53207.
37. Hu H, Tjaden A, Knapp S, Antolin AA, Müller S. A machine learning and live-cell imaging tool kit uncovers small molecules induced phospholipidosis. Cell chemical biology. 2023; 30(12):1634-1651.
38. Geraci J, Rao P, Grandinetti C, Qorri B, Nadolny P, Ayalew K, *et al*. Current Opportunities for the Integration and Use of Artificial Intelligence and Machine Learning in Clinical Trials: Good Clinical Practice Perspectives. Journal of the Society for Clinical Data Management. 2025; 5(2).
39. Luschi A, Nesi P, Iadanza E. Evidence-based clinical engineering: Health information technology adverse events identification and classification with natural language processing. Heliyon. 2023; 9(11).
40. Rane N, Choudhary S, Rane J. Explainable artificial intelligence (XAI) in healthcare: Interpretable models for clinical decision support, 2023. Available at SSRN 4637897.
41. Teixeira B, Carvalhais L, Pinto T, Vale Z. Explainable AI framework for reliable and transparent automated energy management in buildings. Energy and Buildings, 2025, 116246.
42. Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P, *et al*. Explainable AI (XAI): Core ideas, techniques, and solutions. ACM computing surveys. 2023; 55(9):1-33.
43. Nannini L, Balayn A, Smith AL, editors. Explainability in AI policies: A critical review of communications, reports, regulations, and standards in the EU, US, and

UK. Proceedings of the 2023 ACM conference on fairness, accountability, and transparency, 2023.
44. Hadia R. Ethical Principles and Dilemmas in Pharmacy Practice. Chief Editor Prof(Dr) Chhavi Singla. 2024; 41:57.
45. Lu H, Alhaskawi A, Dong Y, Zou X, Zhou H, Ezzi SHA, *et al*. Patient autonomy in medical education: Navigating ethical challenges in the age of artificial intelligence. INQUIRY: The Journal of Health Care Organization, Provision, and Financing. 2024; 61:469580241266364.
46. Fahim YA, Hasani IW, Kabba S, Ragab WM. Artificial intelligence in healthcare and medicine: Clinical applications, therapeutic advances, and future perspectives. European Journal of Medical Research. 2025; 30(1):848.
47. Bourne PA. Evaluating the Quality, Validity, and Trustworthiness of Artificial Intelligence: A Multidisciplinary Synthesis of Secondary Evidence and Governance Frameworks, 2025.
48. Gayathri P. Legal and Ethical dimensions of Artificial Intelligence in Medical Decision-Making, 2025.
49. Shah B, Gupta EV. Ethics and Privacy in AI-Driven Healthcare Decision Support Systems, 2025.
50. Saha K, Okmen N. Artificial Intelligence in Pharmacovigilance: Leadership for Ethical AI Integration and Human-AI Collaboration in the Pharmaceutical Industry, 2025.
51. Mosqueira-Rey E, Hernández-Pereira E, Alonso-Ríos D, Bobes-Bascarán J, Fernández-Leal Á. Human-in-the-loop machine learning: A state of the art. Artificial Intelligence Review. 2023; 56(4):3005-3054.
52. Sayles J. Principles of AI Governance and Model Risk Management: Master the Techniques for Ethical and Transparent AI Systems: Springer Nature, 2024.
53. Kumar R, Yadav A. Navigating Ethical Approval in Legal Research: Global Perspectives and Indian Regulations, 2025.
54. Heidelberg M, Kelman A, Hopkins J, Allen M. The evolution of data ethics in clinical research and drug development. Ethics, Medicine and Public Health. 2020; 14:100517.
55. Nair S. Explainable AI in Gxp Validation: Balancing Automation, Traceability, And Regulatory Trust in The Pharmaceutical Industry. Clinical Medicine and Health Research Journal. 2025; 5(5):1430-1442.
56. Goktas P, Grzybowski A. Shaping the future of healthcare: Ethical clinical challenges and pathways to trustworthy AI. Journal of Clinical Medicine. 2025; 14(5):1605.
57. Daram S. Explainable AI in Healthcare: Enhancing Trust, Transparency, and Ethical Compliance in Medical AI Systems. International Journal of AI, BigData, Computational and Management Studies. 2025; 6(2):11-20.
58. Pasas-Farmer S, Jain R. From discovery to delivery: Governance of AI in the pharmaceutical industry. Green Analytical Chemistry. 2025; 13:100268.