Available online at https://ejournal.corespub.com/index.php/ijhms/index

International Journal of Health, Medicine, and Sports

e-ISSN: 3026-7455

Vol. 3, No. 2, pp. 49-53, 2025

The Role of Explainable Artificial Intelligence (XAI) in Drug Discovery: A Study of Opportunities and Barriers to Implementation

Riza Ibrahim[1*], Hilda Azkiyah[2]

[1]*Research Collaboration Community, Bandung, Indonesia*
[2]*Indonesian Operations Research Association, Bandung, Indonesia*

**Corresponding author email: riza240399@gmail.com*

Abstract

Drug discovery is a complex, lengthy, and costly process with a high failure rate, especially during clinical trials. The integration of Artificial Intelligence (AI) has revolutionized various stages of drug discovery by enabling faster and more accurate analysis of biological and chemical data. However, most AI models in this field operate as "black boxes," where their decision-making processes are opaque and difficult to interpret. This lack of transparency poses significant challenges in terms of trust, validation, and adoption of AI-generated predictions in both clinical and regulatory settings. To address this issue, Explainable Artificial Intelligence (XAI) has emerged as a promising approach to improve the interpretability of AI models without compromising their predictive power. This study aims to systematically review the opportunities, challenges, and future directions of XAI implementation in drug discovery. Using a qualitative method with a systematic literature review approach, data were collected from reputable databases including Scopus, PubMed, IEEE Xplore, SpringerLink, and ScienceDirect, focusing on publications from 2018 to 2024. The analysis identified five main themes: the role of XAI in molecular target identification, application of XAI in compound screening and molecular structure optimization, interpretation of drug toxicity predictions, challenges in XAI implementation, and future research directions. XAI techniques such as SHAP and LIME have proven useful in explaining AI model predictions, improving biological validation, and enabling more informed decision-making by scientists. However, significant challenges remain, including the trade-off between interpretability and accuracy, lack of universal standards, and the complexity of modeling biological systems. This study highlights the critical need for developing standardized interpretability frameworks, user-friendly interfaces, and collaborative environments between data scientists and healthcare professionals to foster XAI adoption in real-world drug discovery processes. Ultimately, XAI has the potential to increase transparency, trust, and efficiency, paving the way for safer and more effective therapeutic developments.

Keywords: Artificial intelligence, explainable artificial intelligence, drug discovery, interpretability, machine learning

1. Introduction

Drug discovery is a complex and time-consuming process, involving the identification of active compounds with therapeutic potential for various diseases (Aware et al., 2022). Traditionally, this process is multi-year and costly, with a high failure rate at the clinical trial stage. Advances in Artificial Intelligence (AI) have brought about a significant change in this approach, enabling more efficient and accurate analysis of biological and chemical data (Bhardwaj et al., 2022).

AI has been used to accelerate various stages of drug discovery, from molecular target identification to compound toxicity prediction (Gupta et al., 2021). However, many of these AI models operate as "black boxes" where the decision-making process cannot be explained transparently (Hassija et al., 2024). This poses challenges in terms of trustworthiness, validation, and adoption of AI models in clinical practice and regulation.

To address these challenges, the Explainable Artificial Intelligence (XAI) approach has emerged, which aims to make the output of AI models more understandable to humans. XAI enables researchers and healthcare professionals to understand the rationale behind model predictions, increasing trust and enabling better validation of the results provided (Yang, 2022; Albahri et al., 2023). The application of XAI in drug discovery brings a variety of opportunities, such as increasing efficiency in compound screening, identifying new biomarkers, and optimizing molecular structures. For example, companies such as Exscientia and BenevolentAI have used AI to accelerate the drug discovery process, with Exscientia being able to reduce development time from years to months (Narayanan et al., 2022).

However, the implementation of XAI also faces several obstacles. One is the trade-off between accuracy and interpretability; models that are easier to explain are often less accurate than complex models such as deep learning. Furthermore, the lack of universal standards in the application of XAI in pharmaceuticals makes it difficult to compare results between studies and to achieve widespread adoption in the industry (Terranova et al., 2024).

Other challenges include the high complexity of biological systems, where complex molecular interactions and biological pathways are difficult to model and explain simply (Cohen et al., 2022). Furthermore, data privacy and security issues are major concerns, given the use of sensitive data such as genomic information and medical records in training AI models.

Although there has been a lot of research in this area, there is still a research gap related to the integration of XAI into existing drug discovery pipelines. Most studies focus on developing accurate AI models, but pay little attention to the aspect of interpretability and how model results can be used practically by scientists and healthcare professionals.

Furthermore, there is still little research evaluating the effectiveness of XAI in real-world contexts, such as how the interpretation of AI models influences clinical decisions or regulatory processes. These studies are essential to understand the true impact of XAI in drug discovery and development. The lack of open and accessible datasets for training and evaluating XAI models is also a barrier to research. The available data are often limited, not standardized, or do not include enough information to allow meaningful interpretation of AI models. Therefore, further research is needed to develop effective XAI methods that can be well integrated into the drug discovery process. This includes developing interpretability standards, creating appropriate datasets, and evaluating the impact of XAI in practical contexts. Thus, XAI can play a significant role in improving efficiency, transparency, and trust in drug discovery and development.

2. Literature Review

Identification of molecular targets is a crucial first step in the drug discovery process. XAI plays a key role in increasing the transparency of the AI models used to identify these targets. By leveraging techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), researchers can understand the contribution of each feature to the model prediction, allowing for better biological validation of the proposed targets (Blanco-Gonzalez, et al., 2023). As in the study by Gimeno et al. (2022), XAI methods were used to link drug screening results to genetic events in acute myeloid leukemia, resulting in predictable and validated therapeutic strategies.

Compound screening and molecular structure optimization are critical stages in drug development. XAI can help explain the reasons behind the selection of certain compounds by the AI model, such as the chemical substructures that contribute to biological activity. This allows chemists to understand and modify the molecular structure more effectively. As an illustration, companies such as Exscientia have successfully reduced the drug development time from years to months by leveraging AI and XAI in this process (Narayanan et al., 2022).

Toxicity prediction is a critical aspect in drug development to ensure the safety of drug candidates. XAI allows for better interpretation of toxicity predictions by AI models, by identifying features that contribute to toxic potential. In a study by Pramudito et al. (2024), the use of XAI with the SHAP method successfully identified optimal in-silico biomarkers for cardiac toxicity evaluation, improving prediction accuracy and providing deeper insight into toxicity mechanisms.

Although XAI offers many benefits, its implementation in drug discovery faces several challenges. One of the main challenges is the trade-off between interpretability and accuracy; models that are easier to explain are often less accurate than complex models such as deep learning. Furthermore, the lack of universal standards in XAI implementation makes it difficult to compare results across studies and for widespread adoption in industry. The complexity of biological systems and the integration of multi-omics data also add to the challenges in developing effective and reliable XAI models (Alizadehsani et al., 2024).

To maximize the potential of XAI in drug discovery, future research needs to focus on developing XAI methods that can be seamlessly integrated into existing drug discovery pipelines. This includes developing interpretability standards, creating appropriate datasets, and evaluating the impact of XAI in practical contexts. In addition, developing intuitive and collaborative user interfaces between data scientists and healthcare professionals will accelerate the adoption of XAI in clinical and regulatory practice (Gangwal & Lavecchia, 2024).

3. Methods

This study uses a qualitative method with a systematic literature review approach. The aim is to analyze and evaluate in depth the opportunities and challenges of implementing Explainable Artificial Intelligence (XAI) in the drug discovery process. This approach was chosen because it allows researchers to comprehensively review various primary and secondary sources, both from scientific journals, industry reports, and official publications of relevant research institutions. The main focus is on the integration of XAI at every stage of drug discovery, from molecular target identification to toxicity evaluation and clinical adoption.

The data used in this study comes from reputable scientific sources, such as Scopus, PubMed, IEEE Xplore, SpringerLink, and ScienceDirect journals. The articles reviewed were selected based on primary keywords such as

"Explainable AI", "Drug Discovery", "XAI in Pharmacology", "AI interpretability", and "SHAP/LIME in drug design". Researchers used a systematic search strategy technique with inclusion criteria such as: articles published between 2018 and 2024, available in English, and explicitly discussing the application of XAI in one of the stages of drug discovery. Exclusion criteria included articles that only discussed AI without aspects of XAI, or that were not relevant to the pharmaceutical and biomedical world.

Data were analyzed using thematic analysis techniques, where each article was reviewed to find the main themes that appeared repeatedly, such as the benefits of XAI in target identification, its role in compound optimization, and the challenges of its implementation. Researchers categorized based on the functional and technical aspects of the application of XAI, then grouped the findings into five main themes as discussed in the literature review. The analysis also included an evaluation of the XAI methods used, such as SHAP, LIME, Grad-CAM, and counterfactual explanation techniques, as well as the context of their use in the studies analyzed.

To ensure the validity and reliability of the study results, researchers used the source triangulation method by comparing results from various types of publications (academic articles, industry reports, white papers, and technical documentation). In addition, content validity is strengthened through a peer debriefing process involving two experts in the fields of AI and pharmacy as external reviewers. The researcher also adopted the PRISMA framework (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) in the literature selection process to be systematic and transparent.

This study has several limitations that need to be considered. First, this study only focuses on literature available online and in English, so the potential for information from local or non-English sources is not covered. Second, because it is qualitative and literature-based, this study does not include empirical tests or direct experiments on the implementation of XAI in drug discovery practices. However, the results of this study are expected to provide a strong theoretical foundation for further, more applicable research in the future.

4. Results and Discussion

4.1. Application Of XAI in Molecular Target Identification

Identification of molecular targets is a crucial first step in the drug discovery process. XAI plays a critical role in increasing the transparency of the AI models used to identify these targets. By leveraging techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), researchers can understand the contribution of each feature to the model prediction, allowing for better biological validation of the proposed targets. For example, in a study by Gimeno et al. (2022), XAI methods were used to link drug screening results to genetic events in acute myeloid leukemia, resulting in predictable and validated therapeutic strategies.

4.2. Molecular Structure Optimization and Compound Screening

Compound screening and molecular structure optimization are critical stages in drug development. XAI can help explain the reasons behind the selection of certain compounds by the AI model, such as the chemical substructures that contribute to biological activity. This allows chemists to understand and modify the molecular structure more effectively. As an illustration, companies such as Exscientia have successfully reduced the drug development time from years to months by leveraging AI and XAI in this process (Narayanan et al., 2022).

4.3. Prediction and Explanation of Drug Toxicity Using XAI

Toxicity prediction is an important aspect in drug development to ensure the safety of drug candidates. XAI allows for better interpretation of toxicity predictions by AI models, by identifying features that contribute to toxic potential. In a study by Pramudito et al. (2024), the use of XAI with the SHAP method successfully identified optimal in-silico biomarkers for cardiac toxicity evaluation, improving prediction accuracy and providing deeper insight into toxicity mechanisms.

4.4. Challenges of XAI Implementation in Drug Discovery

Although XAI offers many benefits, its implementation in drug discovery faces several challenges. One of the main challenges is the trade-off between interpretability and accuracy; models that are easier to explain are often less accurate than complex models such as deep learning. In addition, the lack of universal standards in XAI implementation makes it difficult to compare results between studies and for widespread adoption in the industry. The complexity of biological systems and the integration of multi-omics data also add to the challenges in developing effective and reliable XAI models (Alizadehsani et al., 2024).

4.5. Future Research Directions in XAI Integration and Drug Discovery

To maximize the potential of XAI in drug discovery, future research needs to focus on developing XAI methods that can be seamlessly integrated into existing drug discovery pipelines. This includes developing interpretability standards, generating appropriate datasets, and evaluating the impact of XAI in practical contexts. Additionally, developing intuitive and collaborative user interfaces between data scientists and healthcare professionals will accelerate the adoption of XAI in clinical and regulatory practice (Gangwal & Lavecchia, 2024).

5. Conclusion

The implementation of Explainable Artificial Intelligence (XAI) in drug discovery represents a transformative innovation in modern pharmacology. This study has comprehensively reviewed the role, opportunities, and challenges of XAI across various stages of drug discovery, including molecular target identification, compound screening and optimization, toxicity prediction, and its integration into clinical and regulatory frameworks. The findings highlight that XAI enables better transparency and interpretability of AI model outputs, which are critical for building trust among researchers, clinicians, and regulators. By employing methods such as SHAP, LIME, and other explainability tools, stakeholders gain insights into the rationale behind AI-driven predictions, thereby improving biological validation, optimizing molecular structures, and enhancing the safety profile of drug candidates.

However, despite its promising potential, the implementation of XAI in drug discovery is not without challenges. One prominent issue is the inherent trade-off between model interpretability and predictive accuracy, where highly interpretable models tend to sacrifice performance compared to complex models such as deep learning. Additionally, the absence of standardized frameworks and evaluation metrics for XAI applications in pharmacology limits comparability and hinders broad adoption across the pharmaceutical industry. The complexity of biological systems, multi-omics data integration, and concerns over data privacy and security further complicate XAI implementation.

Given these findings, future research must prioritize the development of standardized interpretability metrics and guidelines to facilitate the practical adoption of XAI in drug discovery pipelines. There is also a need for collaborative platforms that bridge the gap between data scientists and biomedical professionals, enabling more intuitive use of XAI tools in real-world contexts. Moreover, evaluating the real-world impact of XAI, particularly its influence on clinical decision-making and regulatory acceptance, remains an underexplored yet essential area for further investigation. In conclusion, integrating XAI effectively into drug discovery processes holds significant promise for enhancing efficiency, transparency, and trustworthiness, ultimately accelerating the delivery of safer and more effective therapeutics.

References

Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion, 96*, 156-191.

Alizadehsani, R., Oyelere, S. S., Hussain, S., Jagatheesaperumal, S. K., Calixto, R. R., Rahouti, M., ... & De Albuquerque, V. H. C. (2024). Explainable artificial intelligence for drug discovery and development-a comprehensive survey. *IEEE Access*.

Aware, C. B., Patil, D. N., Suryawanshi, S. S., Mali, P. R., Rane, M. R., Gurav, R. G., & Jadhav, J. P. (2022). Natural bioactive products as promising therapeutics: A review of natural product-based drug development. *South African Journal of Botany, 151*, 512-528.

Bhardwaj, A., Kishore, S., & Pandey, D. K. (2022). Artificial intelligence in biological sciences. *Life, 12*(9), 1430.

Blanco-Gonzalez, A., Cabezon, A., Seco-Gonzalez, A., Conde-Torres, D., Antelo-Riveiro, P., Pineiro, A., & Garcia-Fandino, R. (2023). The role of AI in drug discovery: challenges, opportunities, and strategies. *Pharmaceuticals, 16*(6), 891.

Cohen, A. A., Ferrucci, L., Fülöp, T., Gravel, D., Hao, N., Kriete, A., ... & Varadhan, R. (2022). A complex systems approach to aging biology. *Nature aging, 2*(7), 580-591.

Gangwal, A., & Lavecchia, A. (2024). Unlocking the potential of generative AI in drug discovery. *Drug Discovery Today*, 103992.

Gimeno, M., San José-Enériz, E., Villar, S., Agirre, X., Prosper, F., Rubio, A., & Carazo, F. (2022). Explainable artificial intelligence for precision medicine in acute myeloid leukemia. *Frontiers in Immunology, 13*, 977358.

Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Molecular diversity, 25*, 1315-1360.

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., ... & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*, *16*(1), 45-74.

Narayanan, R. R., Durga, N., & Nagalakshmi, S. (2022). Impact of artificial intelligence (AI) on drug discovery and product development. *Indian J. Pharm. Educ. Res*, *56*, S387-S397.

Pramudito, M. A., Fuadah, Y. N., Qauli, A. I., Marcellinus, A., & Lim, K. M. (2024). Explainable artificial intelligence (XAI) to find optimal in-silico biomarkers for cardiac drug toxicity evaluation. *Scientific Reports*, *14*(1), 24045.

Terranova, N., Renard, D., Shahin, M. H., Menon, S., Cao, Y., Hop, C. E., ... & Lu, J. (2024). Artificial intelligence for quantitative modeling in drug discovery and development: an innovation and quality consortium perspective on use cases and best practices. *Clinical Pharmacology & Therapeutics*, *115*(4), 658-672.

Yang, C. C. (2022). Explainable artificial intelligence for predictive modeling in healthcare. *Journal of healthcare informatics research*, *6*(2), 228-239.