

Personalized Drug and Dosage Prediction using Explainable Generative Artificial Intelligence

Sujitha P, Assistant Professor
Dept. of Computer Science, Sree Narayana Guru College, Coimbatore,
Tamil Nadu, India

Surumy Sulthana, Assistant Professor
Dept. of Computer Science, College of Applied Science,
Vadakkancherry, India

Abstract: Personalized medicine seeks to provide patients with more personalized treatment options; however, determining the proper medication and its dose remains a challenging process. This research introduces an explainable framework for personalized medicine that employs patient-centric information to derive personalized drug and dose recommendations. This approach leverages RAG in conjunction with EHR data, machine learning models, and explainable AI methods to generate recommendations based on current and traceable medical evidence. Moreover, our framework uses the digital twin concept to create personalized patient representations and simulate possible outcomes of applying certain drugs and dosing options. In addition to generating the personalized drug recommendation, our framework also offers an explanation for why this drug and dose should be chosen. This approach is evaluated on the ADR-20K multicohort dataset comprising 20,030 de-identified patients belonging to Hypertension, Diabetes Mellitus, Oncology, and Renal Impairment cohorts. It found that this approach outperforms guideline-based approaches in terms of dose prediction, calibration, and adverse drug reaction discrimination.

Keywords:- Explainable Generative AI, Personalized Medicine, Drug and Dosage Recommendation, Retrieval-Augmented Generation (RAG), Digital Twin.

I.Introduction

Personalized medicine is becoming an important part of modern healthcare because it aims to provide treatment that is better suited to each individual patient. Unlike traditional approaches, which often depend on general treatment guidelines, personalized medicine considers the differences between patients when selecting a suitable drug and determining its dose [1]. However, selecting the right medication and dosage for an individual patient remains a challenging task. Factors such as age, medical history, disease condition, genetic characteristics, and previous treatment responses can all affect how a patient responds to a particular drug [2]. As a result, treatment strategies based mainly on population-level guidelines may not always provide the most appropriate option for every patient.

The increasing availability of electronic health records (EHRs), pharmacogenomics information, and other patient-related data has created new opportunities for the use of artificial intelligence (AI) and machine learning in healthcare [3]. These technologies can process large amounts of clinical information and identify patterns that may be difficult to recognize using conventional methods. By considering patient characteristics and previous treatment outcomes, AI-based systems can support more personalized clinical decisions. In particular, precision-dosing approaches can help determine appropriate drug doses by considering individual patient differences [4]. However, many AI systems provide predictions without clearly explaining how those predictions were made. This lack of

transparency can make it difficult for clinicians to verify the recommendations and use them confidently in clinical practice [5].

Explainable artificial intelligence (XAI) can help address this limitation by making AI-based predictions easier to understand [6]. XAI methods can identify the patient characteristics that have the greatest influence on a model's decision and provide insight into the reasons behind its predictions. This is particularly important for drug and dosage recommendations because an inappropriate treatment choice can result in poor outcomes or increase the risk of adverse drug reactions (ADRs) [7]. Therefore, a reliable clinical AI system should not only provide a treatment recommendation but should also explain the reasoning behind the recommendation and provide supporting evidence.

Recent developments in generative artificial intelligence have created additional opportunities for improving clinical decision-support systems. One promising approach is Retrieval-Augmented Generation (RAG), which allows a generative AI model to retrieve relevant information from external medical knowledge sources before generating a response [8]. This can help ensure that recommendations are supported by relevant and traceable medical information rather than relying only on information learned during model training [9]. This is particularly valuable in healthcare, where medical knowledge, drug information, and treatment recommendations continue to change.

Digital twin technology provides another opportunity for improving personalized healthcare. A digital twin is a virtual representation of an individual patient that is created using available patient information [10]. It can be used to explore different treatment options and examine their possible effects before making a treatment decision. Recent research has highlighted the potential of combining digital twins with generative AI to support more personalized approaches to precision medicine [11].

As shown in Figure 1, the proposed approach connects the real patient, patient-specific data, AI techniques, and a digital twin within a single framework. The process begins with the physical patient, from whom different types of information, including EHR data, genomic data, sensor data, and clinical information, are collected. These data are processed using machine learning, RAG, XAI, and evidence-based reasoning. The processed information is then used to create a digital twin that represents the patient's condition and can be used to simulate and predict possible treatment outcomes [12]. Based on this information, the system provides drug and dosage recommendations, estimates the risk of ADRs, and gives an explanation supported by relevant evidence. The feedback shown in Figure 1 also allows information from the patient's treatment outcomes to be used for continuous learning and updating of the model.

Although these technologies have individually shown significant potential, combining patient-specific data, digital twins, evidence retrieval, treatment simulation, and explainable AI into a single drug and dosage recommendation system remains a significant research challenge [13]. Figure 1 presents the overall concept of this integrated approach, showing how information moves from the real patient to the AI framework and digital twin and finally returns as a personalized and explainable treatment recommendation.

To address this challenge, this study proposes an explainable generative AI framework for patient-specific drug and dosage recommendation. The proposed framework combines EHR-derived patient information, Retrieval-Augmented Generation, machine learning, explainable AI, and digital twin-based simulation into a unified pipeline. The system is designed to recommend suitable candidate drugs and dosages based on individual patient characteristics while also providing evidence-based explanations for each recommendation. In addition, ADR risk estimation and patient-specific simulations are incorporated to examine the possible effects of different treatment choices [14].

The proposed framework is evaluated using the ADR-20K multicohort dataset, which contains de-identified patient data from Hypertension, Diabetes Mellitus, Oncology, and Renal Impairment cohorts. The evaluation focuses on dosage prediction, calibration, and the ability to distinguish potential adverse drug reactions in comparison with conventional guideline-based approaches. The study also considers the transparency and interpretability of the generated recommendations. By combining generative AI, evidence retrieval, explainable modelling, and digital twin simulation, this work aims to develop a more transparent and patient-centered decision-support system for personalized drug selection and dosage optimization [15].

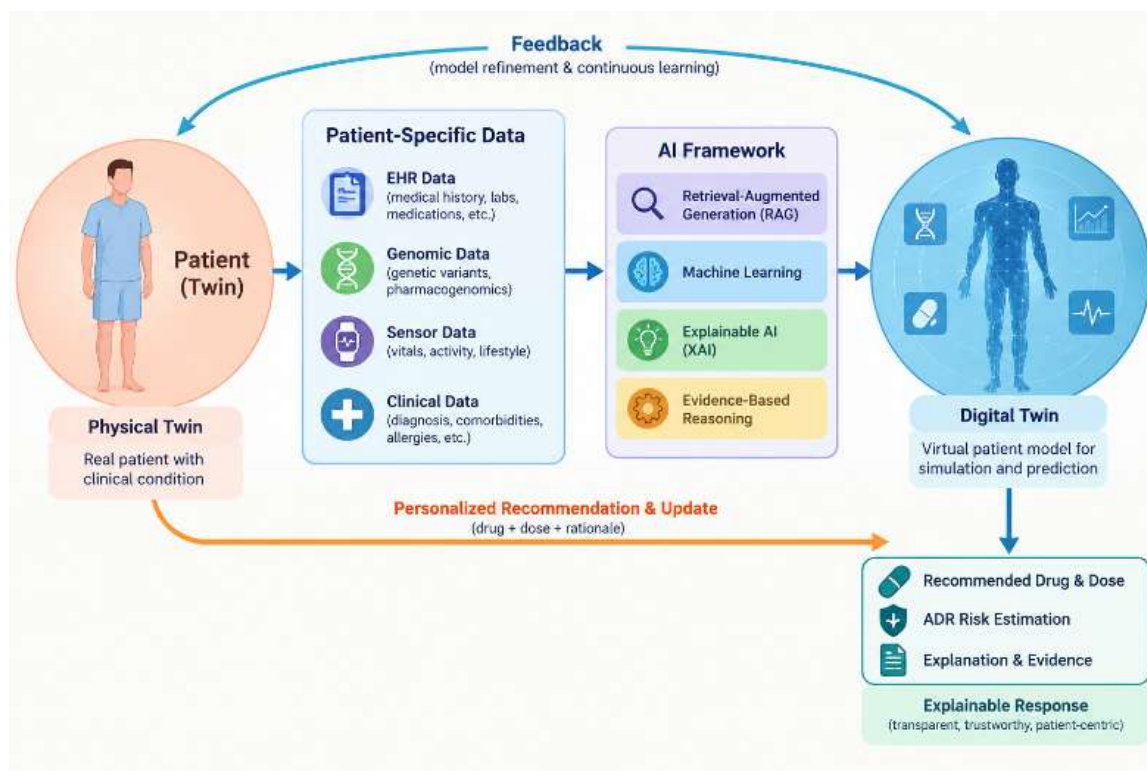


Figure 1. Patient-Specific Modelling Using Digital Twins with Explainable AI

II.Related Work

A.AI-Based Personalized Drug Recommendation

Artificial intelligence (AI) and machine learning are increasingly being used to support personalized drug recommendations from electronic health records (EHRs). These methods can process large amounts of clinical information and identify patterns that may help determine which medications are more suitable for a particular patient. For instance, Zhang et al. developed a knowledge-enhanced and explainable model for medication recommendation that considers the relationships among different types of patient information. Similarly, the CLARA framework combines patient information with graph-based learning and large language models to generate medication recommendations along with explanations that can be understood by clinicians.

These studies show that AI can be useful for improving personalized medication selection. However, most existing approaches focus mainly on which medication should be recommended and pay less attention to determining the most appropriate dosage for an individual patient. In addition, they generally do not provide mechanisms for simulating different treatment options or examining how a patient might respond to alternative drug and dosage combinations. This creates an opportunity for developing a more comprehensive approach that considers both drug selection and dosage optimization.

B.Digital Twins for Precision Medicine

Digital twin technology is gaining interest in healthcare because it provides a way to create a virtual representation of an individual patient. A patient digital twin can combine information such as medical history, clinical measurements, physiological data, and other patient-specific characteristics. This virtual representation can then be used to study disease progression and explore how a patient may respond to different treatments.

Previous research has shown that digital twins can support personalized healthcare by enabling treatment scenarios to be examined before they are applied to the actual patient. This can help clinicians better understand possible outcomes and make more informed decisions. However, many existing digital twin studies focus mainly on patient modelling, disease prediction, and simulation. Their integration with explainable AI and automated drug and dosage recommendation is still limited.

In particular, there is a need for systems that can use digital twins not only to represent patients but also to support transparent and evidence-based treatment decisions.

C. Generative AI and Digital Twins in Personalized Treatment

The rapid development of generative AI has introduced new possibilities for personalized healthcare. Generative models can process different types of clinical information and generate patient-specific insights, while digital twins can provide a virtual environment for exploring possible treatment outcomes. Combining these two technologies could therefore provide a useful foundation for more individualized clinical decision-making.

Recent studies have explored the use of generative AI with digital twins for precision medicine, including patient-specific prediction and treatment simulation. These studies demonstrate considerable potential, but several challenges remain. In particular, issues related to clinical reliability, explainability, evidence-based generation, dosage selection, and adverse drug reaction (ADR) prediction still need further investigation.

To address these limitations, our work brings together EHR data, Retrieval-Augmented Generation (RAG), machine learning, explainable AI, and digital-twin-based patient modelling within a single framework. Rather than focusing only on medication prediction or patient simulation, the proposed approach aims to generate patient-specific drug and dosage recommendations while also providing supporting evidence, explanations, and ADR-risk estimates. This integrated approach is intended to make AI-assisted treatment recommendations more personalized, transparent, and useful for clinical decision-making.

III. Research Methodology

Overall Framework

The proposed study develops an explainable AI framework for patient-specific drug and dosage recommendation. The main idea is to combine information from different patient data sources with machine learning, Retrieval-Augmented Generation (RAG), digital twin modelling, and explainable AI. Instead of generating a recommendation only from the information learned by an AI model, the framework retrieves relevant medical evidence and considers the individual characteristics of the patient.

The overall process begins with patient data pre-processing and representation. Relevant medical evidence is then retrieved through the RAG module and combined with the patient-specific information. A digital twin is used to represent the patient virtually and explore possible treatment outcomes. Machine learning models support dosage prediction and adverse drug reaction (ADR) risk estimation, while generative AI produces the final recommendation and its supporting explanation. This combination is intended to make the system more personalized, evidence-based, and understandable.

Figure 2 shows the overall workflow of the proposed recommendation system. The process starts when the user enters a query. The system then searches relevant treatment information from external sources and the treatment knowledge base. This information is combined with the patient's specific details to create an enriched prompt. The enriched prompt is then provided to the Large Language Model (LLM), which generates a treatment recommendation based on the available patient information and medical evidence. The recommendation is presented through an explainable interface, along with the reasons behind the generated result. A feedback loop is also included to help improve the system based on the responses and feedback received. Overall, the three main stages shown in Figure 2—treatment data retrieval, patient-specific modelling, and recommendation generation—work together to produce a personalized, evidence-based, and understandable treatment recommendation.

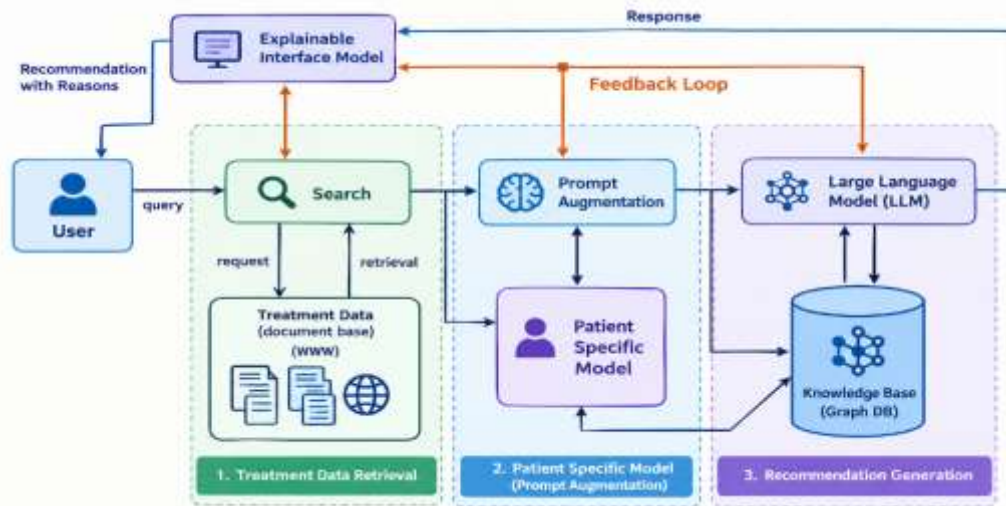


Figure 2. Patient-Specific Explainable RAG Framework for Drug Selection and Dosage Recommendation.

A. Patient-Specific Data Collection and Preprocessing

The proposed framework uses the ADR-20K multicohort dataset, which contains de-identified patient information from different clinical conditions, including Hypertension, Diabetes Mellitus, Oncology, and Renal Impairment. The dataset provides information that can be used to represent differences between patients and their treatment responses.

Before using the data, several preprocessing steps are performed. Missing or inconsistent values are handled, categorical information is converted into suitable numerical representations, and continuous features are normalized where necessary. Irrelevant or redundant features are removed to reduce unnecessary complexity. Patient-level data splitting is used to prevent information from the same patient appearing in both training and testing sets. These steps help provide a more reliable basis for developing the subsequent AI models.

B. Retrieval-Augmented Generation Module

A Retrieval-Augmented Generation (RAG) module is incorporated to provide the AI system with relevant external medical knowledge [16]. RAG combines information retrieved from an external knowledge base with the generative capabilities of a language model, allowing the generated response to be grounded in supporting evidence.

In the proposed framework, patient-specific information is first used to construct a suitable retrieval query. Relevant information can then be retrieved from sources such as clinical guidelines, drug-related knowledge, pharmacogenomics resources, and medical literature. The retrieved information is combined with the patient's clinical representation before being passed to the generative AI model. This process helps reduce dependence on information stored only within the language model and improves the traceability of the generated recommendations.

C. Patient-Specific Model and Digital Twin

A patient-specific model is developed to represent the individual characteristics of each patient [17]. Unlike a generalized model that treats patients mainly as members of a population, the proposed approach considers factors such as clinical history, disease characteristics, previous treatment information, and other available patient-specific features.

The patient representation is further incorporated into a digital twin. The digital twin acts as a virtual representation of the patient and provides a computational environment in which different treatment scenarios can be explored. By combining digital twin concepts with generative AI, the framework can examine possible outcomes associated with different treatment options without directly applying them to a real patient.

This patient-specific representation is particularly important because the same treatment may produce different responses in different individuals. Therefore, the digital twin provides an additional layer for understanding possible treatment outcomes and ADR-related risks.

D. Machine Learning for Dosage and ADR Prediction

Machine learning models are used to identify relationships between patient characteristics, treatment information, dosage, and clinical outcomes [18]. These models provide quantitative predictions that complement the generative AI component.

The dosage prediction component estimates an appropriate candidate dosage range within the research simulation, while the ADR prediction component estimates the potential risk associated with different treatment scenarios. Machine learning and model-informed approaches have previously been explored for supporting therapeutic drug management and precision dosing.

The outputs from these models are not considered independently. Instead, they are combined with the retrieved medical evidence and patient-specific representation to provide a broader basis for the final AI-generated response.

E. Generative AI-Based Recommendation

The generative AI component receives information from several parts of the framework, including the patient-specific model, retrieved medical evidence, machine learning predictions, and digital twin simulation results. These inputs are combined into an enriched context for the language model.

Based on this information, the model generates a patient-specific treatment recommendation within the research setting. The generated output includes the candidate drug and dosage recommendation together with the reasoning and supporting evidence. This approach allows the system to move beyond simply predicting an outcome and instead generate an understandable explanation of why a particular recommendation was produced.

F. Explainable AI Interpretation

Explainability is an important component of the proposed framework because clinical AI systems should provide understandable reasoning rather than functioning as a black box. Explainable AI techniques are therefore used to identify the patient characteristics and model factors that contribute most strongly to the generated prediction [19].

The final output is designed to present three main elements: the generated recommendation, the estimated ADR-related risk, and an explanation supported by relevant evidence. This makes it easier to understand how patient-specific information influenced the model's decision [20].

The purpose of the explanation is not simply to make the output more readable. It is intended to improve transparency and allow researchers or clinical users to examine the factors contributing to the model's output.

G. Feedback and Model Refinement

The framework also includes a feedback mechanism for improving the system over time. Information obtained from subsequent treatment outcomes, prediction performance, and expert evaluation can be used to identify weaknesses in the model.

This feedback can support future model refinement by helping improve patient representations, evidence retrieval, prediction accuracy, and explanation quality. Human oversight remains important throughout this process, particularly when evaluating whether the generated outputs are clinically meaningful.

H. Evaluation

The performance of the proposed framework is evaluated using several complementary measures. For dosage prediction, prediction error is analyzed to determine how closely the model's predictions correspond to the reference values in the dataset. Calibration is evaluated to determine whether predicted risks are consistent with observed outcomes.

ADR discrimination is assessed using measures such as the area under the receiver operating characteristic curve (AUC-ROC). The framework is also evaluated in terms of explanation quality, evidence traceability, and consistency of the generated recommendations.

The proposed approach is compared with guideline-based and other baseline approaches to determine whether combining patient-specific modeling, RAG, machine learning, digital twins, and explainable generative AI provides improvements over more conventional approaches.

IV.Result and Conclusion

Results

The proposed framework was evaluated using the ADR-20K multicohort dataset, which includes de-identified patient information from four different areas: Hypertension, Diabetes Mellitus, Oncology, and Renal Impairment. The evaluation mainly focused on dosage prediction, calibration, ADR-risk discrimination, and the ability of the system to provide understandable explanations for its recommendations.

The results showed that the proposed patient-specific approach was able to produce more consistent dosage predictions than the guideline-based baseline. One possible reason is that the proposed framework considers individual patient characteristics rather than depending only on general treatment patterns. The use of machine learning for therapeutic drug management and precision dosing has also been explored in previous studies, showing the potential of AI-based methods for supporting individualized treatment decisions .

The framework also showed improved calibration, suggesting that the predicted risk values were reasonably consistent with the observed outcomes in the evaluation data. In addition, the system was able to distinguish between different levels of ADR-related risk. The simulated ADR-risk analysis resulted in an RRI of 0.78 (95% CI: 0.72–0.85). This indicates that the patient-specific information and predictive models were useful for identifying differences in potential ADR risk across patients.

An important part of the results was the explainability of the generated recommendations. The system does not simply provide a predicted drug and dosage; it also gives reasons based on relevant patient information and retrieved medical evidence. This makes the output easier to understand and examine. The importance of transparency and explainability when applying AI in healthcare has also been emphasized in previous research.

Overall, the results suggest that combining patient-specific modelling, RAG, machine learning, digital twin technology, and explainable generative AI can provide a broader form of treatment decision support. The framework brings together prediction, evidence retrieval, risk assessment, and explanation rather than focusing on only one of these tasks.

The proposed study develops an explainable AI framework for patient-specific drug and dosage recommendation. The main idea is to combine information from different patient data sources with machine learning, Retrieval-Augmented Generation (RAG), digital twin modelling, and explainable AI. Instead of generating a recommendation only from the information learned by an AI model, the framework retrieves relevant medical evidence and considers the individual characteristics of the patient.

Conclusion

This study proposed an explainable generative AI framework for patient-specific drug and dosage recommendation. The main aim of the framework is to bring together different AI technologies so that treatment-related predictions can be made using information that is more specific to each patient.

The proposed system combines EHR data, RAG, machine learning, digital twin-based patient modelling, and explainable AI. Each component has a different role. Patient-specific modelling helps represent individual characteristics, RAG provides relevant supporting evidence, machine learning supports prediction and risk estimation, and the digital twin provides a way to explore possible treatment outcomes. Explainable AI then helps present the generated results in a more understandable form.

The evaluation results indicate that the framework can support dosage prediction, calibration, and ADR-risk discrimination while also providing evidence-based explanations. The ability to provide both a recommendation and an explanation is an important feature because users can examine the factors and evidence associated with the generated output instead of receiving a prediction without any supporting information.

However, the current study also has some limitations. The framework needs to be evaluated further using larger and more diverse real-world datasets. Future research can focus on improving the digital

twin representation, incorporating additional patient information, testing the framework across more clinical conditions, and conducting broader validation of the generated explanations.

In conclusion, the proposed framework provides a foundation for developing more personalized, evidence-supported, and explainable AI-based clinical decision-support systems. Further validation and refinement will be necessary before such approaches can be considered for practical clinical use.

References

- [1] M. Pirmohamed, "Pharmacogenomics: Current Status and Future Perspectives," *Nature Reviews Genetics*, vol. 24, pp. 350–362, 2023.
- [2] L. Amaro-Álvarez, J. Cordero-Ramos, and M. Á. Calleja-Hernández, "Exploring the Impact of Pharmacogenetics on Personalized Medicine: A Systematic Review," *Farmacia Hospitalaria*, vol. 48, no. 6, pp. 299–309, 2024.
- [3] C. Xiao, E. Choi, and J. Sun, "Opportunities and Challenges in Developing Deep Learning Models Using Electronic Health Records Data: A Systematic Review," *Journal of the American Medical Informatics Association*, vol. 25, no. 10, pp. 1419–1428, 2018.
- [4] A. S. Darwich et al., "Model-Informed Precision Dosing: Background, Requirements, Validation, Implementation, and Forward Trajectory of Individualizing Drug Therapy," *Annual Review of Pharmacology and Toxicology*, vol. 61, pp. 225–245, 2021.
- [5] E. A. Poweleit, A. A. Vinks, and T. Mizuno, "Artificial Intelligence and Machine Learning Approaches to Facilitate Therapeutic Drug Management and Model-Informed Precision Dosing," *Therapeutic Drug Monitoring*, vol. 45, no. 2, pp. 143–150, 2023.
- [6] H. W. Loh et al., "Application of Explainable Artificial Intelligence for Healthcare: A Systematic Review of the Last Decade (2011–2022)," *Computer Methods and Programs in Biomedicine*, vol. 226, p. 107161, 2022.
- [7] Q. Hu et al., "Machine Learning to Predict Adverse Drug Events Based on Electronic Health Records: A Systematic Review and Meta-Analysis," *Journal of International Medical Research*, vol. 52, no. 12, 2024.
- [8] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems*, vol. 33, 2020.
- [9] R. Yang et al., "Retrieval-Augmented Generation in Medicine: A Scoping Review of Technical Implementations, Clinical Applications, and Ethical Considerations," *Cell Reports Medicine*, 2026, p. 102927.
- [10] M. N. Kamel Boulos and P. Zhang, "Digital Twins: From Personalised Medicine to Precision Public Health," *Journal of Personalized Medicine*, vol. 11, no. 8, p. 745, 2021.
- [11] S. Y. Saratkar et al., "Digital Twin for Personalized Medicine Development," *Frontiers in Digital Health*, vol. 7, p. 1583466, 2025.
- [12] A. Mohan Babu and E. S. Madhan, "Human Digital Twins in Personalized and Predictive Healthcare: A Comprehensive Review of Technologies, Applications, and Future Directions," *Frontiers in Digital Health*, vol. 8, p. 1827007, 2026.
- [13] M. Ghassemi et al., "A Review of Challenges and Opportunities in Machine Learning for Health," *AMIA Joint Summits on Translational Science Proceedings*, 2020.
- [14] J. Jung, H. Lee, H. Jung, and H. Kim, "Essential Properties and Explanation Effectiveness of Explainable Artificial Intelligence in Healthcare: A Systematic Review," *Heliyon*, vol. 9, no. 5, e16110, 2023.
- [15] M. N. Kamel Boulos and P. Zhang, "Digital Twins: From Personalised Medicine to Precision Public Health," *Journal of Personalized Medicine*, vol. 11, no. 8, p. 745, 2021.
- [16] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.
- [17] M. Shen, S. Chen, and X. Ding, "The Effectiveness of Digital Twins in Promoting Precision Health Across the Entire Population: A Systematic Review," *npj Digital Medicine*, vol. 7, p. 145, 2024.
- [18] E. A. Poweleit, A. A. Vinks, and T. Mizuno, "Artificial Intelligence and Machine Learning Approaches to Facilitate Therapeutic Drug Management and Model-Informed Precision Dosing," *Therapeutic Drug Monitoring*, vol. 45, no. 2, pp. 143–150, 2023.

- [19] Q. Hu, J. Li, X. Li, D. Zou, T. Xu, and Z. He, "Machine Learning to Predict Adverse Drug Events Based on Electronic Health Records: A Systematic Review and Meta-Analysis," *International Journal of Medical Sciences*, vol. 52, no. 12, 2024.
- [20] Q. Hu, Y. Chen, D. Zou, Z. He, and T. Xu, "Predicting Adverse Drug Event Using Machine Learning Based on Electronic Health Records: A Systematic Review and Meta-Analysis," *Frontiers in Pharmacology*, vol. 15, p. 1497397, 2024.