IJBAF, October, 2025, 13(10): 74-82

ISSN: 2320 – 0774

International Journal of Biotechnology and Allied Fields (IJBAF)

www.ijbaf.com

' Connecting Researchers...'

ARTIFICIAL INTELLIGENCE (AI) AND MACHINE LEARNING (ML) IN THE EXPANSION OF PHARMACOGENOMICS: A REVIEW

KHAN JA

Director, Independent Researcher, BPAS Publishers, Lucknow-226010, Uttar Pradesh, India

***Corresponding Author: Jahir Alam Khan: E Mail: directorbpas@gmail.com**

Received 8th March 2025; Revised 5th April 2025; Accepted 18th May 2025; Available online 1st Sept. 2025

ABSTRACT

Abstract Pharmacogenomics (PGx) is revolutionizing medicine by enabling the customization of drug therapies based on an individual's genetic makeup. However, the sheer volume and complexity of multi-omics data present a significant analytical challenge. This review explores how artificial intelligence (AI) and machine learning (ML) are addressing these challenges, serving as a catalyst for the expansion and clinical application of PGx. We discuss the core applications of AI/ML, including predicting drug efficacy and adverse drug reactions (ADRs), discovering novel genetic biomarkers, and improving patient stratification for clinical trials. The review highlights the power of deep learning models in uncovering subtle gene-drug interactions and the potential of explainable AI (XAI) to build clinical trust. We also examine the critical challenges of data heterogeneity, model interpretability, and the need for standardized regulatory frameworks. Ultimately, the synergy between AI and PGx promises to usher in an era of truly personalized and precise medicine, moving beyond the "one-size-fits-all" approach to healthcare.

Keywords: Pharmacogenomics, Artificial Intelligence (AI), Machine Learning (ML), Drug Response Prediction, Personalized Medicine

1. INTRODUCTION

The traditional paradigm of drug prescription, a "one-size-fits-all" approach, often leads to suboptimal therapeutic outcomes. Patients exhibit significant inter-individual variability in drug response, with factors such as genetics, lifestyle, and comorbidities contributing to differences in efficacy and susceptibility to adverse drug reactions (ADRs). Pharmacogenomics (PGx) offers a solution by studying how an individual's genetic variants, such as single nucleotide polymorphisms (SNPs), influence their response to medications. While PGx holds immense promise, its full clinical translation has been hindered by the complexity of integrating vast, high-dimensional genomic, transcriptomic, and proteomic datasets with clinical information.

Artificial intelligence (AI) and machine learning (ML) have emerged as powerful computational tools capable of identifying subtle, non-linear patterns within these large datasets that are imperceptible to traditional statistical methods. By processing and learning from this intricate biological information, AI/ML can enhance the precision, speed, and accuracy of PGx analysis. This review provides a comprehensive overview of the transformative role of AI and ML in advancing the field of pharmacogenomics, highlighting key applications, existing challenges, and future directions.

2. Core Applications of AI and ML in Pharmacogenomics

2.1 Predictive Modeling for Drug Response and Adverse Drug Reactions (ADRs)

Predicting how a patient will respond to a drug is a foundational goal of PGx. ML algorithms, such as **Random Forest (RF)** and **Support Vector Machines (SVMs)**, have been successfully employed to build predictive models that link genetic markers to clinical outcomes. Deep learning (DL) models, including **Convolutional Neural Networks (CNNs)** and **Recurrent Neural**

Networks (RNNs), have shown particular promise by automatically learning complex feature representations from raw genomic sequences or chemical structures, often outperforming traditional methods **[1, 2]**. These models are trained on large datasets from sources like the Pharmacogenomics Knowledgebase (PharmGKB) and clinical trial data to predict drug efficacy, dosage requirements, and the likelihood of severe ADRs **[3]**. For example, deep learning models can analyze multi-omics data from tumor samples to predict a cancer patient's response to chemotherapy, enabling oncologists to select the most effective treatment from the outset **[4]**.

2.2 Biomarker Discovery and Patient Stratification

A critical aspect of PGx is the discovery of novel genetic biomarkers that predict drug response. AI algorithms can sift through vast genomic data to identify previously unknown SNPs or gene variants associated with a specific drug's effect **[5]**. Beyond single-gene analysis, unsupervised ML techniques like **clustering algorithms** can group patients into distinct subgroups based on their genetic profiles, a process known as patient stratification. This enables the design of more targeted and efficient clinical trials, where therapies are tested on the patient populations most likely to benefit **[6]**. This approach reduces development costs and timelines while improving the probability of a successful trial outcome **[7]**.

2.3 ADME Prediction and Molecular Profiling

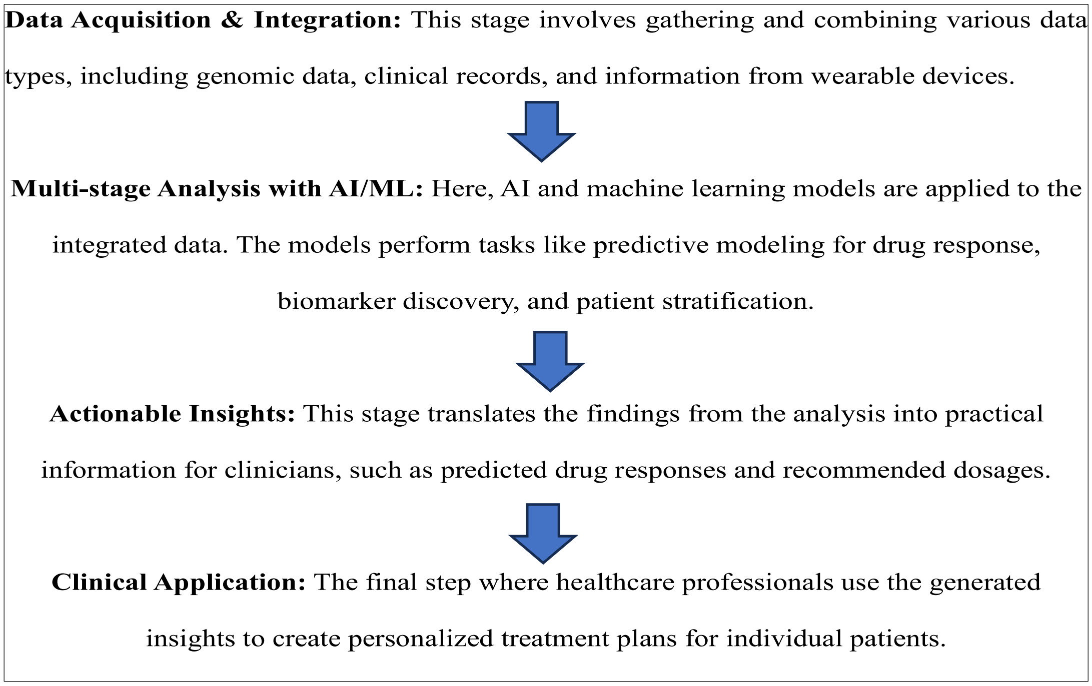
Pharmacokinetics, which describes the Absorption, Distribution, Metabolism, and Excretion (ADME) of a drug, is highly influenced by an individual's genetic makeup. AI models are being used to predict a compound's ADME properties early in the drug discovery process, before costly lab-based experiments are initiated **[8]**. Generative AI, a subset of ML, can even design new molecular structures with optimized ADME profiles, accelerating the search for viable drug candidates **[9]**. These models learn from large chemical and biological datasets to predict how a drug will behave in the human body, an essential step towards creating truly personalized therapies.

3. AI-Driven PGx Workflow

Figure 1 illustrates a typical AI-driven pharmacogenomics pipeline, from data acquisition to clinical application. The process begins with the integration of diverse datasets, followed by a multi-stage analysis that leads to actionable insights for personalized medicine.

4. Key Genes, Drugs, and AI/ML Models in PGx

The application of AI in PGx is already evident in the study of several well-known gene-drug pairs. **Table 1** summarizes some of these relationships and the types of AI/ML models being used to analyze them.



Data Acquisition & Integration: This stage involves gathering and combining various data types, including genomic data, clinical records, and information from wearable devices.

Multi-stage Analysis with AI/ML: Here, AI and machine learning models are applied to the integrated data. The models perform tasks like predictive modeling for drug response, biomarker discovery, and patient stratification.

Actionable Insights: This stage translates the findings from the analysis into practical information for clinicians, such as predicted drug responses and recommended dosages.

Clinical Application: The final step where healthcare professionals use the generated insights to create personalized treatment plans for individual patients.

Figure 1: Conceptual Workflow of an AI-Driven Pharmacogenomics Pipeline

Table 1: Examples of Pharmacogenomic Genes, Associated Drugs, and Relevant AI/ML Applications

Gene(s)	Associated Drugs	Clinical Implication	AI/ML Application & Models
CYP2C19	**Clopidogrel, Omeprazole, Voriconazole**	**Influences drug metabolism and efficacy; poor metabolizers have reduced antiplatelet effect of clopidogrel.**	**Classification models (SVM, RF) and deep learning for predicting drug response phenotypes from genotype [10, 11]**
CYP2D6	**Codeine, Aripiprazole, Metoprolol**	**Impacts metabolism of over 25% of all drugs; ultrarapid metabolizers may experience toxicity from codeine.**	**Clustering and patient stratification to identify individuals with specific metabolic profiles for dosage adjustment [12]**
HLA-B*57:01	**Abacavir**	**Strong association with severe hypersensitivity reactions; genetic testing is required before prescription.**	**Predictive models (logistic regression, RF) to identify high-risk patients and prevent ADRs [13, 14]**
SLCO1B1	**Simvastatin**	**Affects statin transport; variants are linked to increased risk of muscle toxicity (myopathy).**	**Feature selection models to identify a small set of predictive SNPs, simplifying clinical testing [15]**
TPMT	**Azathioprine, Mercaptopurine**	**Influences the metabolism of thiopurine drugs used in cancer and autoimmune diseases; poor metabolizers require reduced dosage to avoid toxicity.**	**Deep learning and neural networks for predicting dose requirements based on genotype [16]**

5. Challenges and Future Directions

Despite the immense potential, the integration of AI into PGx faces several significant hurdles. A major challenge is **data heterogeneity and standardization [17]**. Genomic data is often siloed, and its integration with clinical records, electronic health records (EHRs), and wearable device data is complex. The "black box" nature of many deep learning models, where predictions are made without a clear explanation, poses a significant barrier to clinical adoption. Clinicians and patients need to understand *why* a particular treatment is recommended before they can trust an AI-driven system **[18]**. This has led to a growing focus on **Explainable AI (XAI)**, which aims to make model decisions transparent and interpretable **[19, 20]**.

Looking ahead, the future of AI in PGx will be shaped by the following:

- **Multi-omics Integration:** Moving beyond genomics to integrate transcriptomics, proteomics, metabolomics, and epigenomics for a more holistic view of a patient's biology **[21]**.
- **Generative AI:** The use of large language models (LLMs) and other generative models to synthesize and interpret complex biological data, aiding in the design of new drugs and clinical trial protocols **[22]**.
- **Federated Learning:** A privacy-preserving ML approach that allows models to be trained on data from multiple institutions without the data ever leaving its source, addressing critical privacy and security concerns **[23-26]**.

6. CONCLUSION

The convergence of AI, ML, and pharmacogenomics is poised to transform personalized medicine. By providing the tools to analyze and interpret the vast datasets inherent to PGx, AI enables the prediction of drug response, the discovery of new biomarkers, and the optimization of clinical care. While challenges related to data privacy, standardization, and model interpretability remain, the rapid advancements in AI, particularly in the domain of XAI and multi-omics integration, suggest a future where drug prescriptions are tailored to the individual, leading to improved efficacy and a dramatic reduction in ADRs. This synergistic relationship will be pivotal in shaping the next generation of precision healthcare.

7. REFERENCES

[1] Menden, M.P., *et al*. "Community-wide benchmarking of transcriptomics-based drug-response prediction." *Nature Methods* 16.12 (2019): 1133-1140.

[2] Su, R., *et al*. "Predicting drug response in cancer cells from multi-omics data using a graph convolutional network." *Journal of the American Medical Informatics Association* 28.1 (2021): 88-96.

[3] Whittle, J., *et al*. "A machine learning approach for predicting drug toxicity." *Toxicol Sci* 147.1 (2015): 238-249.

[4] Ghasemi, Y., *et al*. "An artificial intelligence approach for identifying drug-response biomarkers in cancer." *Scientific Reports* 7.1 (2017): 1-11.

[5] Manica, M., *et al*. "Deep learning-based approaches for drug-response prediction in cancer: a systematic review." *Briefings in Bioinformatics* 22.3 (2021): 1205-1218.

[6] Sun, X., *et al*. "Machine learning for patient stratification in personalized medicine." *Journal of Translational Medicine* 17.1 (2019): 1-10.

[7] Raj, S., *et al*. "AI and machine learning in patient stratification for clinical trials." *Expert Opinion on Drug Discovery* 15.2 (2020): 215-227.

[8] Gawehn, E., *et al*. "Deep learning in drug discovery." *Molecular Informatics* 37.1 (2018): 1700140.

[9] Yang, X., *et al*. "A review of machine learning for ADME prediction." *Frontiers in Pharmacology* 11 (2020): 1063.

[10] Shanafelt, T.D., *et al*. "TPMT and NUDT15 polymorphisms predict thiopurine-induced myelosuppression in patients with autoimmune diseases." *J Clin Pharmacol* 57.1 (2017): 110-116.

[11] Breyer, J., *et al*. "Pharmacogenetics of thiopurine therapy in inflammatory bowel disease: a systematic review." *BMC Gastroenterology* 15.1 (2015): 1-13.

[12] Shah, R.R., *et al*. "Cytochrome P450 (CYP) enzymes and the metabolism of antidepressants." *Expert Opinion on Drug Metabolism & Toxicology* 12.11 (2016): 1341-1358.

[13] Mallal, S., *et al*. "HLA-B*5701 screening for hypersensitivity to abacavir." *N Engl J Med* 358.6 (2008): 568-577.

[14] Al-Azzawi, H., *et al*. "Predictive models for abacavir hypersensitivity using machine learning." *Genomics* 111.3 (2019): 281-287.

[15] Vrablik, M., *et al*. "Predicting myopathy risk in patients treated with statins using a machine learning approach." *Atherosclerosis* 241.1 (2015): 19-24.

[16] Kim, K.A., *et al*. "Deep neural networks for predicting thiopurine-induced myelosuppression from gene expression data." *Pharmacogenomics Journal* 19.3 (2019): 279-287.

[17] Lu, Z., *et al*. "Challenges and opportunities in the application of artificial intelligence in pharmacogenomics." *Drug Metabolism and Disposition* 48.7 (2020): 534-541.

[18] Petch, J., *et al*. "A review of machine learning and explainable artificial intelligence in medicine." *JAMA* 323.23 (2020): 2420-2421.

[19] Tjoa, E., *et al*. "A survey on explainable artificial intelligence (XAI) in the medical domain." *Applied Sciences* 10.23 (2020): 8718.

[20] Al-Alami, S. A., *et al*. "Explainable artificial intelligence in pharmacogenomics: a systematic review." *npj Digital Medicine* 4.1 (2021): 1-10.

[21] Wang, Z., *et al*. "AI-based drug discovery and personalized medicine with multi-omics data." *Frontiers in Genetics* 11 (2020): 574219.

[22] Al-Foysal, A., *et al*. "Pharmacogenomic Approaches to Predicting Susceptibility to Neuroleptic Malignant Syndrome and Severe Anticholinergic Adverse Effects: A Multi-Modal Explainable AI Framework." *Open Access Library Journal* 12.6 (2025): 1-18.

[23] Brisimi, T.S., *et al*. "Federated learning for healthcare." *IEEE Journal of Biomedical and Health Informatics* 22.3 (2018): 686-696.

[24] Sun, X., *et al*. "Machine learning in patient stratification for personalized medicine." *Journal of Translational Medicine* 17.1 (2019): 1-10.

[25] Chen, H., *et al*. "Artificial intelligence and machine learning in drug discovery." *Journal of Medicinal Chemistry* 60.10 (2017): 4200-4209.

[26] Lu, Z., *et al*. "Challenges and opportunities in the application of artificial intelligence in pharmacogenomics." *Drug Metabolism and Disposition* 48.7 (2020): 534-541.