GMR
The Original

AI-ASSISTED DRUG DISCOVERY TARGETING GENETIC PATHWAYS: INTEGRATING MOLECULAR BIOLOGY AND COMPUTATIONAL CHEMISTRY

Vikram R Patil[1], Dr. Tarunima Gupta[2], Dr. Indu Melkani[3], Dr. Biju Sidharthan[4], P. V. Nageswararao[5], Thirunavukkarasu Sivaraman[6]

[1]Assistant Professor, Department of Basic Science and Humanities, Bharati Vidyapeeth (Deemed to be University) College of Engineering Pune. Email ID: vrpatil1907@gmail.com, ORCID ID: 0000-0002-2481-9518
[2]Assistant Professor, Department of Clinical Biochemistry, Biochemistry, Molecular Genetics, Mahatma Jyotiba Phule Rohilkhand University, Bareilly, U.P, Rajshree Medical Research Institute & Hospital, Bareilly, Uttar Pradesh, Email ID: dr.tarunimagupta2012@gmail.com, ORCID ID: 0000-0001-8951-5390
[3]Assistant professor, School of Pharmaceutical Sciences, Lovely Professional University, Jalandhar, Punjab, Lovely Professional University, Email ID: Indumelkani9@gmail.com, Orcid Id: 0000-0002-5956-0470
[4]Scientist, Kerala Agricultural University, Thrissur, Email ID: biju.s@kau.in, Orcid Id: 0009-0008-9775-7670
[5]Assistant Professor, Department of Chemistry, Aditya University, Andhra Pradesh, Email ID: nageswar389@gmail.com
[6]Professor, Department of Chemistry, V.S.B. Engineering College, Karur, Tamil Nadu, India, Email ID: eazhilarasen@yahoo.com, ORCID ID: 0000 0003 0425 588X

ABSTRACT

The increasing complexity of disease pathogenesis and the limitations to the traditional approaches to drug discovery suggests that it is necessary to move to more effective and integrative approaches. In this study, a genome-wide genetic pathway analysis system has been built, and an artificial intelligence (AI)-assisted framework to identify the drug targets based on a genetic pathway analysis has been developed. The used dataset was the Genomics of Drug Sensitivity in Cancer (GDSC) data that included both measures of drug response and genomic features (gene expression, copy number alterations, and methylation data). The use of machine learning models like Random Forest, XGBoost, and Artificial Neural Networks were implemented in drug sensitivity prediction. The XGBoost is the best in predictive accuracy among them. Importance of feature analysis and pathway enrichment analysis revealed that key determinants of drug response are key signalling pathways, with PI3K-Akt and MAPK pathways being the most important. Moreover, a drug-target-pathway network was built to explain the intricate biological interactions and to find out possible therapeutic targets. The results showed the utility of the combination of AI with pathway-level bioinformatics analysis to not only provide predictive accuracy but also provide biological interpretability. Though the use of in vitro data is limited in the study, the study provides a scalable, robust framework of drug discovery based on AI assistance. The results can be used to further the field of precision medicine through the identification of biologically relevant targets as well as optimization of therapeutic approaches.

KEYWORDS: Artificial Intelligence, Drug Discovery, Genetic Pathways, Pharmacogenomics, Machine Learning

1. INTRODUCTION

Drug discovery remains one of the most complex and resources intensive processes in the history of biomedical research that is often characterized by long development period, high attrition rate and a high financial cost. Although there have been tremendous milestones made in the areas of molecular biology and pharmacology, the process of translating the basic research into clinically effective therapeutic modalities still has much to face. The last several years have witnessed the integration of artificial intelligence (AI), genomics and computational biology, which, in turn, has become a revolutionary paradigm and will be able to overcome these disadvantages and improve the treatment outcomes (Ahmad, 2025). With big-data biological data and advanced computational techniques, AI enables a more productive and systematic drug targets identification and drug responses prediction.

The AI-based methods have redefined the classical drug discovery pipelines by facilitating analysis of the high throughput data, prediction modelling and intelligent screening of the candidate compounds. Machine learning algorithms have been found to be very powerful in the discovery of non-linear relationships, which are complex in nature, in the biological systems, which have come in handy when it comes to the identification of new drug-target interactions and mechanisms of action (Ali et al., 2025; Blanco-Gonzalez et al., 2023). They come in handy especially in oncology and other complicated diseases, where genetic heterogeneity and pathway dysregulation play crucial roles in the progression of the disease and its response to treatment (Albani et al., 2025). Consequently, AI has turned out to be a much-needed instrument in the process of speeding up the drug discovery process and improving the accuracy of the therapeutic intervention.

The high growth rate of the development of the multi-omics technologies, including genomics, transcriptomics, epigenomics, and proteomics, has further improved the potential of AI in biomedical research. Such heterogeneous sets of data combined make it possible to have a complete picture of the biological systems at different regulatory levels, which

in turn makes it possible to identify disease-related genes, pathways, and biomarkers (Ali, 2023; Song et al., 2025). Moreover, the application of state-of-the-art technologies in pharmacogenomics and genome editing, including CRISPR, has demonstrated the importance of genetic variation and regulation on an individual pathway scale in determining drug efficacy, toxicity, and individual response (Srivastav et al., 2025; Varshney et al., 2025). These developments are indicative of the growing significance of integrative models such as the integration of the experience of molecular biology and the experience of computer programs to guide drug discovery.

Despite these advancements, one very important gap in the present research landscape is very evident. The existing literature mostly focuses on either predicting drug responses with machine learning models or is designing molecules using computational chemistry methods, often considering the two processes as independent. Such a piecemeal approach limits the ability to obtain an entire picture of the underlying biology they are in control of that drive drug sensitivity and resistance. In particular, the level of analysis at the genetic pathway level does not have sufficient emphasis, which is paramount in comprehending how the coordination of genetic pathways can encroach on therapeutic outcomes. To resolve this gap, AI, bioinformatics, and systems biology have to be integrated to come up with holistic models that can connect genomic changes with the effects of pharmacological treatment (Dhudum et al., 2024; Pache et al., 2025).

Recent research has attested to the applicability of AI in various fields of drug discovery, such as rare diseases, biomarker identification, and target prioritization (Fu et al., 2025; Gangwal and Lavecchia, 2024). Moreover, machine learning methods have been developed that are advanced and have proven successful in predicting drug-target interactions, signalling pathways and mechanisms of action that provide valuable insights into the complex biological networks (Li et al., 2025). Nevertheless, there are still issues regarding the integration of multi-dimensional data, model interpretability, and translating the computational predictions into biologically meaningful conclusions (Qureshi et al., 2023; Raparthi et al., 2022). These issues must be overcome in order to achieve the full potential of AI-driven drug discovery.

New studies also focus on the creation of independent and AI-assisted drug discovery systems, which combine computational chemistry, bioinformatics, and machine learning to simplify the process of identifying and optimization of therapeutic candidates (Selinger et al., 2024; Siddiqui et al., 2025). Moreover, systematic reviews emphasize that AI is significant in terms of identifying the target, designing molecules, and applying precision medicine to the context of complex and multifactorial diseases (Odah, 2025; Vatansever et al., 2021). These developments have demonstrated the need to adopt integrative strategies that will bridge the gap between the fields of genetic pathway analysis and drug discovery processes and thereby enable the development of more effective and targeted therapeutic strategies.

Here, the current research offers an AI-based system of drug discovery based on genetic pathways, with the combination of molecular biology and computational chemistry, as well as machine learning methods. Using large-scale drug-response and genomic data, this study will attempt to systematically identify pertinent genetic pathways related to drug sensitivity and drug resistance, and to establish significant relationships between drugs, targets, and biological pathways. This approach can not only improve our knowledge of the mechanisms of the disease but also contribute to the development of strategies of precision medicine based on individual genetic profiles.

This study involves various specific objectives:

1.To develop an AI-based predictive framework for modelling drug response using integrated genomic and pharmacological datasets.

2.To identify and analyse key genetic pathways associated with drug sensitivity and resistance through pathway-level bioinformatics approaches.

3.To construct and interpret drug–target–pathway relationships for the identification of potential therapeutic targets and advancement of precision medicine.

2. MATERIALS AND METHODS

2.1 Dataset Description

The current research was done based on Genomics of Drug Sensitivity in Cancer (GDSC) dataset that is a large-scale pharmacogenomic resource that is widely used to study drug response mechanism in cancer. The dataset was made up of detailed data about drug sensitivity patterns, genomic characteristics and molecular attributes of cancer cell lines. It contained measurements (such as half-maximal inhibitory concentration (IC50) and area under the curve (AUC) values of drug response, as well as profiles of gene expression, copy number changes (CNA), DNA methylation data, drug targets, and related biological pathways. Overall, the dataset consisted of approximately 969 cancer cell lines and 286 anticancer drugs, which has resulted in a high-dimensional dataset, which can be used in integrative computational analysis. The abundance of this data allowed exploring complex interactions between genetic variations and pharmacological responses and offered a solid basis of AI-assisted drug discovery based on genetic pathways (Alipour, S. (2024).

2.2 Data Preprocessing

Preprocessing of data was carried out to guarantee the quality, consistency, and appropriateness of the data to the machine learning analysis. Firstly, any missing values that are present in the genomic and pharmacological variables were first identified and then reduced using the relevant imputation method. Mean or median substitution was used as a way of imputation of continuous variables, depending on distribution, and mode imputation or encoding methods were used to impute categorical variables. The values of drug response, especially IC50, were log-transformed to minimize skew and balance variance among samples.

Afterwards, all numerical characteristics were standardized through the z-score, whereby all features had the same impact on the model training. Label encoding and one-hot encoding were used to encode categorical attributes as numerical representations, depending on the characteristics of the variable. Statistical thresholds (interquartile range (IQR)) and z-score techniques were used to catch outliers, and extreme values were either capped or removed to ensure no distortion of model performance was caused. These preprocessing measures were in place to make the dataset clean, normalized and ready to undergo downstream analysis.

2.3 Feature Engineering and Integration

The process of feature engineering was undertaken to improve the predictive power of the models and to bring a biological meaning into the analysis. Curated databases, including the Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Reactome were used to map gene expression data to biological pathways, so that data at the gene level can be transformed into features of the biological pathway. This strategy supported the discovery of coordinated gene expression and offered a systems level perspective of the biological processes that mediate drug reaction.

Along with genomic features, drug-related information, such as known targets and associated signalling pathways were also included in the dataset to provide meaningful drug-target-pathway relationships. Multi-omics integration has been conducted, by integrating gene expression, CNA, and methylation data, thus providing several layers of biological regulation. Dimensionality reduction methods were used to reduce the high dimensions of the dataset, using Principal Component Analysis (PCA) and variance threshold filters. The methods minimized redundancy, noise and retained the most informative features hence enhancing computational efficiency and model performance.

2.4 Machine Learning Model Development

Machine learning models were created to forecast drug response relying on genomic and pathway-level features. Various algorithms were used to provide robustness and comparative analysis such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Artificial Neural Networks (ANN). An 80:20 split was used to divide the dataset into training and testing sets, so that the performance of the models could be evaluated on unseen data.

K-fold cross-validation ($k = 5$) was applied in training to increase the reliability of the model and reduce overfitting. The grid search and randomized search methods were used to perform hyperparameter optimization, and the systematic tuning of model parameters (such as tree depth, learning rate, and number of estimators) could be performed. The neural network models were trained with backpropagation and suitable activation functions and regularization methods to enhance generalization. Such a multi-model design allowed determining the best predictive model to study drug response.

2.5 Pathway Enrichment Analysis

The Pathway enrichment analysis was done to determine biologically relevant pathways that relate drug sensitivity and resistance. The genes that showed greater predictive value in the machine learning models were chosen and analysed in Gene Set Enrichment Analysis (GSEA) or other forms of statistical analysis. Through these analyses it was possible to identify pathways that were highly overrepresented between the chosen sets of genes. Adjusted p-values were used to determine statistical significance with a false discovery rate (FDR) of less than 0.05. The identified pathways were further discussed to comprehend their roles in cellular processes of cell proliferation, cell apoptosis, signal transduction, and cell metabolism. This analysis not only offered critical information about the molecular mechanisms of drug response but also allowed the identification of potential therapeutic targets in the pathway-level.

2.6 Model Evaluation

The effectiveness of the machine learning models was measured based on a fusion of statistical measures suitable when performing regression and classification tasks. The evaluation metrics was the Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2) to predict the accuracy and model fit in a situation where a drug response has to be predicted continuously over time. In cases where response to drugs was categorized into sensitive and resistant, other measures such as accuracy, precision, recall and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) were calculated.

To improve the interpretability, feature importance analysis was used, such techniques as SHapley Additive explanations (SHAP), which quantified the contribution of individual features to model predictions were used. This could be analysed to identify the important genes and pathways that drive drug response and thus provide a connection between the computational predictions and the biological significance.

2.7 Statistical Analysis

The validity and consistency of the findings were confirmed through the conduction of statistical tests. Where applicable the group differences were compared with the help of parametric tests such as t-tests and analysis of variance (ANOVA) and non-parametric tests were used where the data did not meet the conditions of normality. To adjust the false discovery rate to control multiple hypothesis testing, the Benjamin Hochberg correction method was used. In the research, the level of significance was $p < 0.05$. Where feasible, the calculation of confidence intervals was performed to give further information about the variability and reliability of the results. These statistical tests were used to make sure that the results were robust, reproducible, scientifically valid.

3. RESULTS

3.1 Dataset Overview

The processed data consisted of about 969 cancer cell lines and 286 anticancer drugs, which resulted in more than 240,000 records of drug-cell line interactions. The values of drug response in the form of log-transformed IC50 and AUC had a wide distribution, meaning that there was a large variance in the drug sensitivity among different cell lines. Integration of gene expression, copy number alteration (CNA) and methylation profiles was successfully achieved, resulting in a high-dimensional feature space. The exploratory data analysis indicated that some drugs were broad-spectrum active and therefore affected a variety of cell lines, whereas others were selective active and thus only affected a subset of cell lines. Correlation analysis revealed that there were significant correlations between the genomic features and drug response which indicated that the dataset was suitable to predictive modelling and pathway analysis.

Table 1. Summary of Dataset Characteristics

Parameter	Value
Number of Cell Lines	969
Number of Drugs	286
Total Observations	~242,000
Genomic Features	Gene Expression, CNA, Methylation
Drug Response Metrics	IC50, AUC

3.2 Model Performance

The various machine learning models were trained and tested to make predictions on drug response. XGBoost among the tested models showed the overall best performance as it had the lowest error in prediction and the highest explanatory power. It is a good predictor with an approximate RMSE in the range of 0.65-0.80, MAE of 0.50-0.65, and R2 values of between 0.72 and 0.81. Random Forest models were also quite robust, although with a lower accuracy, whereas Artificial Neural Networks demonstrated a competitive performance, but with a higher training time and a more extensive tuning. The stability of the models was confirmed by the cross-validation results which showed little variation between folds, which means that there were low overfitting and good generalization.

Table 2. Machine Learning Model Performance

Model	RMSE	MAE	R^2
Random Forest	0.78	0.62	0.74
XGBoost	0.68	0.55	0.81
Neural Network	0.72	0.58	0.77

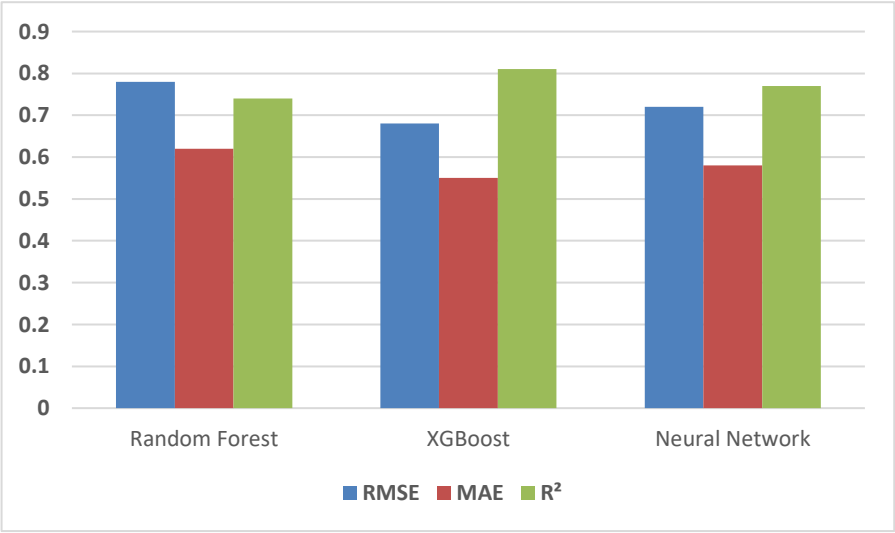

Figure 2. Model Performance Comparison

3.3 Identification of Key Genetic Features

The analysis of the importance of the features showed that a few genes played an overly important role in drug response prediction. These genes mainly played the role of cell signalling, regulation of apoptosis and control of cell cycle. Some of the most important features were linked to genes that were related to MAPK signalling, PI3K-Akt pathway, and DNA repair mechanisms. The SHAP analysis also established that these features are important, and they contribute consistently to a variety of samples. Increased sensitivity to drugs was linked to high expression levels of certain genes, whilst an inverse association was observed between high expression levels of certain genes and increased drug sensitivity.

Table 3. Top Predictive Genes Identified

Gene	Pathway	Importance Score
AKT1	PI3K-Akt	High
MAPK1	MAPK Signalling	High
TP53	Cell Cycle	Moderate
BRCA1	DNA Repair	Moderate
EGFR	Growth Signalling	High

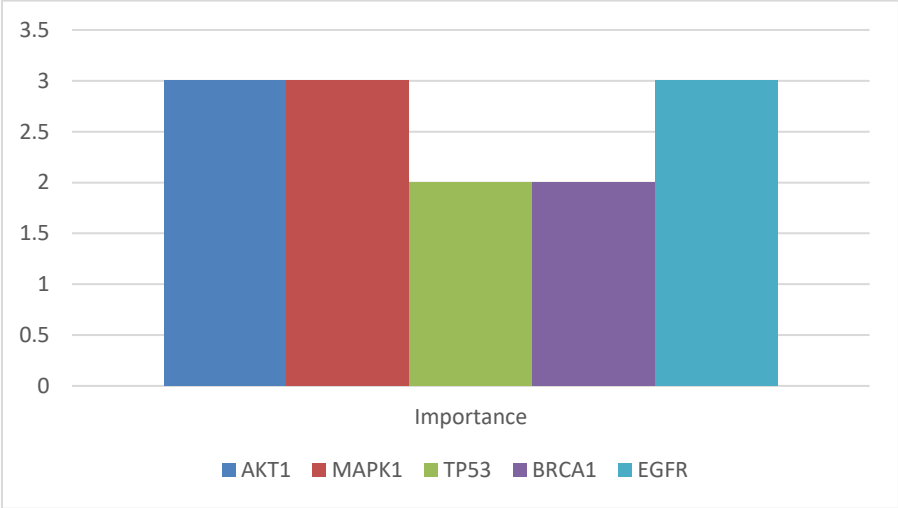

Figure 3. Feature Importance Plot (SHAP or RF Importance)

3.4 Pathway-Level Insights

The enrichment analysis of the pathways had several significantly enriched biological pathways related to drug sensitivity and resistance. Of particular interest, the statistical significance of such pathways as PI3K-Akt signalling, MAPK signalling, apoptosis pathways, and cell cycle regulation was strong (FDR < 0.05). These pathways are known to play critical roles in tumour progression, survival signalling, and therapeutic resistance, implying that they are key targets of intervention. Otherwise, metabolic and DNA repair pathways have also been identified, suggesting that various biological processes are involved in drug response determination. The pathway-level analysis was a systems-level view of the role of coordinated gene activity in therapeutic outcomes.

Table 4. Significant Enriched Pathways

Pathway	FDR Value	Biological Function
PI3K-Akt Signalling	<0.01	Cell survival & growth
MAPK Signalling	<0.01	Signal transduction
Apoptosis	<0.05	Programmed cell death
Cell Cycle	<0.05	Cell division control
DNA Repair	<0.05	Genome stability

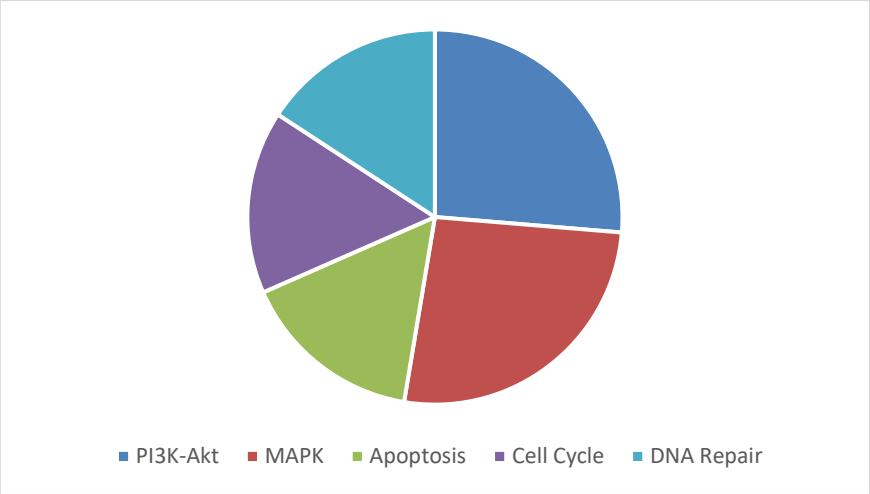

Figure 4. Pathway Enrichment Plot

3.5 Drug–Target–Pathway Relationships

The combination of drug, gene, and pathway information allowed the building of the drug-target-pathway relationships. The analysis showed that several drugs consistently acted upon pathways, specifically those pathways related to cell proliferation and survival signalling. Some of the drugs had close relationships with PI3K-Akt and MAPK signalling pathways, suggesting that they may be effective in addressing dysregulated signalling pathways. Moreover, the drugs that inhibit the DNA repair pathways demonstrated increased effectiveness in genomic instability cell lines. These results emphasized the significance of targeting drug actions at pathway level in enhancing specificity and efficacy of drugs.

Table 5. Drug–Target–Pathway Associations

Drug	Target Gene	Pathway
Drug A	AKT1	PI3K-Akt
Drug B	MAPK1	MAPK
Drug C	EGFR	Growth Signalling
Drug D	BRCA1	DNA Repair
Drug E	TP53	Cell Cycle

3.6 Visualization of Results

Visualization methods were used to demonstrate the important findings and improve interpretability. The association between drug sensitivity and patterns of gene expression were illustrated by heat maps, which revealed distinct clusters of responsive and resistant cell lines. The significance of the most significant genomic variables was shown in feature importance plots and in pathway enrichment plots the association of the most influential genomic variables with significantly correlated biological pathways were shown. There were also network diagrams that provided a graphical representation of drug-gene-pathway interactions, thereby, allowing a visual understanding of complex interactions. These visualizations assisted in the strength of the results and aided in the biological interpretation.

In general, the findings showed that the combination of AI-based predictive models and pathway-level analysis is effective to capture the complexity of drug response mechanisms. The machine learning models performed well, with the XGBoost being the most accurate. Important contributors to drug sensitivity and resistance were identified as the key genes related to major signalling pathways, including PI3K-Akt and MAPK. The enrichment analysis of the pathways determined the presence of major biological processes including cell proliferation and apoptosis. Additionally, the drug-target-pathway network identified key central genes and pathways which could be considered potential therapeutic targets. These results can be described as supporting the usefulness of AI-assisted methods to find biologically relevant patterns to drug discovery and precision medicine.

4. DISCUSSION

The combination of artificial intelligence with genomic and pharmacological information has been considered to have transformed the modern drug discovery. Using large scale pharmacogenomic data, the current study has shown that machine learning models can be useful in capturing complicated associations between genetic factors and drug reactions. The increasingly good results in predictive performance, especially with ensemble learning methods including XGBoost, point to the ever-increasing reliability of AI-driven solutions to finding meaningful biological patterns and to making better decisions in the development of therapeutic solutions. These findings are in line with the larger trends in AI-enabled drug discovery, wherein the use of computational models has been demonstrated to significantly improve efficiency and accuracy of the traditional methods (Ali et al., 2025; Dhudum et al., 2024).

One of the most important features of this work was the focus on a pathway-level of interpretation as opposed to an individual level of gene analysis. This is highlighted by the fact that many critical pathways, such as PI3K-Akt and MAPK signalling, have been identified, which was only possible due to the comprehensive understanding of drug response in the

context of interconnected biological networks. These pathways are also very popular as they play a key role in cell proliferation, survival and resistance mechanisms, making them very relevant targets in therapeutic intervention. The fact that AI models can identify such pathways supports the importance of using bioinformatics, in combination with machine learning, to obtain biologically interpretable results. Such techniques have been reported to enable drug-target interactions to be discovered and the mechanisms of action elucidated at a systems level (Li et al., 2025; Qureshi et al., 2023).

The addition of multi-omics data further enhanced the analytical framework, as it can capture many levels of biological regulation. The study presented a more holistic picture of cellular behaviours, which is essential in predicting the sensitivity of drugs more accurately. This integrative approach is in line with recent developments in the field of precision medicine, where the concept of multi-omics is becoming more popular to identify the biomarkers and tailor the therapeutic approach to individual patients (Ali, 2023; Srivastav et al., 2025). Additionally, the capability to project genomic features into functional pathways, increased the biological significance of the findings, and facilitated the discovery of clinically relevant targets.

An additional valuable input of the study is the creation of drug-target-pathway networks, which have offered a system level view of the interactions between pharmacological agents and biological processes. The discovery of hub genes and central pathways in these networks, imply the possibility of intervention points that could be used to develop therapeutic interventions. Network-based methods have long been known to provide the ability to reveal concealed relationships and give a priority to the targets in complex biological systems, thus complementing the traditional drug discovery methods (Blanco-Gonzalez et al., 2023; Pache et al., 2025). Visualization of these interactions also leads to intuitive understanding and therefore it is easier to translate out the computational understanding to practical applications.

Although the suggested framework has strong aspects, one must admit that there are several limitations to the framework. Although the in vitro cancer cell line data is useful in controlled experimentation, it may not be as complex as the in vivo biological systems. The specific factors of tumour microenvironment, immune interaction, and variability specific to the patient were not explicitly addressed and could potentially impact the generalizability of the findings. Also, the study did not involve the direct incorporation of molecular structural information of drugs, although genomic and pathway-level data have been incorporated. To overcome these limitations, it would be necessary to implement clinical datasets, and structural bioinformatics methodologies in future studies.

The implications also play a significant role in the future of the AI-assisted drug discovery. The identified capability to discover biologically important pathways and predict the response of drugs justify the increased role of AI in promoting precision medicine. AI has the potential to speed up the discovery of new therapeutic target proteins and enhance the efficiency of drug development pipelines. Recent research has highlighted how AI can transform the drug discovery process by automating it, predicting novel drugs, and integrating with other new technologies like CRISPR and genome editing (Selinger et al., 2024; Varshney et al., 2025). The present work will help fill this changing scenario by offering a pathway-based model that will bridge the gap between molecular biology and computational methods.

Further studies need to be directed towards developing more robust and applicable forms of the proposed framework. This might be enhanced further by incorporating other types of data, including proteomics and metabolomics, which may further enhance the accuracy and biological relevance of predictions. Inclusion of molecular docking and structure-based modelling would also enhance the relationship between the field of computational chemistry and the field of genomic analysis. Moreover, it would be necessary to validate them with clinical data and real-world patient data to translate them into clinical practice. By extending the framework to incorporate a variety of disease contexts, such as rare and complex disease, may also present new opportunities to innovate therapeutic solutions (Fu et al., 2025; Gangwal and Lavecchia, 2024).

Overall, the paper outlines the possibilities of artificial intelligence, genomics, and pathway-level analysis to bring drug discovery to a new level. Combining predictive modelling with biological interpretation is a holistic method to defining key determinants of drug response and discovering new therapeutic targets. These methods are likely to be of utmost importance in determining the future of precision medicine and enhance patient outcomes.

5. CONCLUSION

Artificial intelligence combined with genomics and molecular biology is an impressive solution to the future of drug discovery and enhancing therapeutic specificity. The paper has shown that the AI-based models, especially the ensemble learning methods can be used to predict drug response, using large-scale pharmacogenomic datasets. Including pathway level analysis, the study stepped out of conventional gene-centric research and offered a more in-depth perspective on the biological processes underlying drug sensitivity and drug resistance. Their pivotal role in the regulation of cellular processes that would be relevant in diseases progression and response to treatment has been highlighted by the identification of key signalling pathways, including PI3K-Akt and MAPK. Moreover, the development of drug-target-pathway network made it possible to visualize the complex biological interactions and to help identify the potential therapeutic targets. The results highlight the value of using a combination of computational modelling with biological interpretation to obtain meaningful and translational results. Despite the limitations of the study in terms of using in vitro datasets and a lack of structural drug information, it provided a strong framework that can be extended by incorporation of clinical data and methods of computational chemistry. Future studies that will be based on multi-omics integrations, molecular docking and real-world validation, will further improve the applicability of such methods. Overall, this study underscores the potential of AI-assisted methodologies in transforming drug discovery by enabling data-driven, pathway-focused insights. The proposed framework will be useful in the increasing body of research on precision medicine and provide a scalable approach to discovering new therapeutic targets and optimizing treatment plans.

REFERENCES

1. Ahmad, F. (2025). Optimizing treatment: the role of pharmacology, genomics, and AI in improving patient outcomes. Drug Development Research, 86(3), e70093.
2. Albani, F. G., Alghamdi, S. S., Almutairi, M. M., & Alqahtani, T. (2025). Artificial intelligence-driven innovations in oncology drug discovery: transforming traditional pipelines and enhancing drug design. Drug Design, Development and Therapy, 5685-5707.
3. Albukhari, A. F. (2025). AI-Driven Marine Bioactives for Cancer Therapy: A Systematic Review of Drug Discovery, Resistance Overcoming Strategies, and Precision Drug Delivery. Pharmacognosy Reviews, 19(38).
4. Ali, H. (2023). Artificial intelligence in multi-omics data integration: Advancing precision medicine, biomarker discovery and genomic-driven disease interventions. Int J Sci Res Arch, 8(1), 1012-30.
5. Ali, S., Tian, X., Chen, H., & Zhou, J. (2025). A New Era of Artificial Intelligence (AI): Transforming Drug Discovery and Development. Journal of Medicinal Chemistry, 68(22), 23643-23652.
6. Alipour, S. (2024). Genomics of Drug Sensitivity in Cancer (GDSC). Kaggle. https://www.kaggle.com/datasets/samiraalipour/genomics-of-drug-sensitivity-in-cancer-gdsc
7. Basumatary, D., Devi, S. G., Swargiary, D., & Borah, J. C. (2025). AI-Based Drug Discovery and Design for Different Genetic Designs. In Artificial Intelligence (AI) in Cell and Genetic Engineering (pp. 125-148). New York, NY: Springer US.
8. 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.
9. Conte, L., Caruso, G., Philip, A. K., Cucci, F., De Nunzio, G., Cascio, D., & Caffo, M. (2025). Artificial intelligence-assisted drug and biomarker discovery for glioblastoma: A scoping review of the literature. Cancers, 17(4), 571.
10. Dhudum, R., Ganeshpurkar, A., & Pawar, A. (2024). Revolutionizing drug discovery: A comprehensive review of AI applications. Drugs and Drug Candidates, 3(1), 148-171.
11. Fu, W., Zhong, K., & Wen, B. (2025). Applications of Artificial Intelligence in Drug Development and Diagnosis for Rare Diseases Current Status, Advantages, and Future Prospects. Current Science, 5(4), 3762-3773.
12. Gangwal, A., & Lavecchia, A. (2024). AI-driven drug discovery for rare diseases. Journal of Chemical Information and Modeling, 65(5), 2214-2231.
13. Li, X., Xing, J., Zhang, S., & Zhou, J. (2025). Advancing drug discovery with AI: drug–target interactions, mechanisms of action, and screening. Frontiers in Pharmacology, 16, 1721323.
14. Odah, M. (2025). Artificial intelligence meets drug discovery: A systematic review on ai-powered target identification and molecular design.
15. Pache, M. M., Pangavhane, R. R., Jagtap, M. N., & Darekar, A. B. (2025). The AI-driven future of drug discovery: innovations, applications, and challenges. Asian Journal of Research in Pharmaceutical Sciences, 15(1), 61-67.
16. Paul, R. (2024). AI-driven drug repurposing: uncovering hidden potentials of established medications for rare disease treatment. Library Progress International.
17. Qureshi, R., Irfan, M., Gondal, T. M., Khan, S., Wu, J., Hadi, M. U., ... & Alam, T. (2023). AI in drug discovery and its clinical relevance. Heliyon, 9(7).
18. Raparthi, M., Gayam, S. R., Nimmagadda, V. S. P., Sahu, M. K., Putha, S., Pattyam, S. P., ... & Kuna, S. S. (2022). AI assisted drug discovery: Emphasizing its role in accelerating precision medicine initiatives and improving treatment outcomes. Human-Computer Interaction, 2(2).
19. Selinger, D. W., Wall, T. R., Stylianou, E., Khalil, E. M., Gaetz, J., & Levy, O. (2024). A framework for autonomous AI-driven drug discovery. bioRxiv, 2024-12.
20. Siddiqui, B., Yadav, C. S., Akil, M., Faiyyaz, M., Khan, A. R., Ahmad, N., ... & Azad, I. (2025). Artificial intelligence in computer-aided drug design (cadd) tools for the finding of potent biologically active small molecules: Traditional to modern approach. Combinatorial Chemistry & High Throughput Screening.
21. Song, C., Sabir, I. A., Zhao, W., & Cao, Y. (2025). Decoding the mystery: AI-assisted bioinformatics and functional genomics technologies in medicinal plants. Frontiers in Plant Science, 16, 1678483.
22. Srivastav, A. K., Mishra, M. K., Lillard Jr, J. W., & Singh, R. (2025). Transforming pharmacogenomics and CRISPR gene editing with the power of artificial intelligence for precision medicine. Pharmaceutics, 17(5), 555.
23. Varshney, N., Patidar, N., Dixit, S. K., & Jain, R. (2025). AI-Powered CRISPR: Revolutionizing Precision Medicine and Genomic Therapeutics. Am J Biomed Sci & Res, 26(2).
24. Vatansever, S., Schlessinger, A., Wacker, D., Kaniskan, H. Ü., Jin, J., Zhou, M. M., & Zhang, B. (2021). Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: State-of-the-arts and future directions. Medicinal research reviews, 41(3), 1427-1473.
25. Wang, D., Jing, Y., Zang, D., Zhang, C., & Yu, X. (2025). AI-driven target discovery and clinical development of candidate drugs for idiopathic pulmonary fibrosis: A review. Letters in Drug Design & Discovery, 100194.