ACCSCIENCE PUBLISHING

Eurasian Journal of Medicine and Oncology

ORIGINAL RESEARCH ARTICLE

Integrating data analytics into health informatics: Advancing equity, pharmaceutical outcomes, and public health decision-making

Md. Majedur Rahman[1], Md. Shihab Rahman[1], Safiul Islam[1], Sajidul Islam Khan[1], Abdullah Al Mahmud Ashik[1], Emran Hossain[1], and Ahmed Tanvir[2]*

[1]College of Graduate and Professional Studies, Trine University, Angola, Indiana, United States of America

[2]Department of Neurology, School of Medicine, Louisiana State University Health Shreveport, Shreveport, LA, United States of America

Abstract

Introduction: The integration of data analytics into health informatics has become vital for transforming raw clinical information into actionable insights that improve patient care and pharmaceutical outcomes.

Objectives: This study uses the Medical Information Mart for Intensive Care IV (MIMIC-IV) electronic health record dataset to examine differences in pharmaceutical prescription patterns and their relationship to clinical outcomes. We investigate how demographic characteristics, including age, gender, and race, affect prescribing patterns for three major drug classes: opioids, antibiotics, and antipsychotics.

Methods: We analyzed the MIMIC-IV intensive care unit dataset, incorporating preprocessing of demographic and prescription data to support fairness and outcome analysis. A decision tree model was trained to predict in-hospital mortality and evaluated using standard performance metrics.

Results: We examined the relationship between drug type and patient outcomes, finding that antibiotic prescriptions were associated with shorter hospital stays, whereas antipsychotic prescriptions were linked to longer hospitalizations. Our findings reveal statistically significant differences in prescribing patterns, where men were more likely to receive opioids, whereas women were more likely to receive antibiotics. In addition, considerable racial disparities suggest possible systemic inequities. Nevertheless, there was no statistically significant correlation between drug type and in-hospital mortality, indicating that underlying clinical conditions may play a more substantial role. The model achieved an area under the receiver operating characteristic curve of 0.9337 and an F1-score of 0.8235, outperforming several complex algorithms whereas remaining easily interpretable—an important advantage in clinical practice.

Conclusion: These results demonstrate the potential of transparent machine learning models to support enhanced medical decision-making and highlight the need for prescription strategies that prioritize fairness and equity.

Keywords: Clinical outcome prediction; Data analytics; Health informatics; Machine learning; Prescription equity

***Corresponding author:**
Ahmed Tanvir
(tanvir.ahmed@lsuhs.edu)

Citation: Rahman MM, Rahman MS, Islam S, *et al*. Integrating data analytics into health informatics: Advancing equity, pharmaceutical outcomes, and public health decision-making. *Eurasian J Med Oncol*. 2025;9(4):284-295. doi: 10.36922/EJMO025300319

Received: July 23, 2025

Revised: September 15, 2025

Accepted: September 25, 2025

Published online: November 7, 2025

1. Introduction

Integrating data analytics into health informatics is increasingly essential for enhancing clinical treatments, optimizing pharmacotherapy, and addressing health disparities. The widespread availability of electronic health records (EHRs) has enabled the use of large-scale real-world clinical data to evaluate treatment effectiveness, uncover inequities, and inform evidence-based health policy.[1] This transformation has significantly influenced both clinical decision-making and the design of public health studies. Bhuiyan and Mondal[2] and Abbas *et al*.[3] emphasized that artificial intelligence (AI)-driven predictive analytics in health care can significantly improve cost efficiency and operational effectiveness by leveraging large-scale clinical data.

In critical care settings, the impact of different medication categories (e.g., antimicrobials, vasopressors, and sedatives) on patient outcomes such as mortality, length of stay (LOS), and complication risks can be substantial.[4] Thus, it is crucial to understand not only the effects of medications but also the variability in patient responses across different populations and demographic variables. Existing literature indicates that both systemic injustices and biological differences contribute to variations in pharmacotherapy outcomes; however, these issues are often underexplored in traditional clinical trials.[5]

Medical Information Mart for Intensive Care IV (MIMIC-IV) provides a comprehensive and de-identified dataset comprising clinical information from over 70,000 intensive care unit (ICU) patients. This resource allows researchers to investigate treatment patterns and outcomes across diverse patient cohorts. For instance, machine learning methods applied to this dataset can model patient outcomes, such as in-hospital mortality, based on specific administered drugs. These models can also be evaluated across sensitive demographic categories such as gender and race.[6]

Moreover, without deliberate efforts to monitor and address fairness, there is growing concern that predictive algorithms in health care may perpetuate or even exacerbate pre-existing disparities. Evidence suggests that models trained on unrepresentative or imbalanced datasets can lead to algorithmic bias and unequal care delivery.[7] Therefore, the implementation of machine learning in health informatics must include responsible monitoring, assessment, and evaluation of fairness.

Furthermore, privacy-preserving data analytics is gaining importance. Recent reviews on federated learning in smart health care have shown that sensitive data can be analyzed collaboratively across institutions without centralizing records.[8] Similarly, HeartEnsembleNet, a hybrid ensemble model for cardiovascular risk prediction, demonstrates how ensemble methods can maintain interpretability while achieving high performance.[9] These advances complement our emphasis on model transparency and suggest future directions for integrating equity-focused data analytics with advanced machine learning approaches.[10,11]

In this study, we analyzed the MIMIC-IV database with a focus on drugs typically administered in ICU settings, particularly opioids and antibiotics, to explore potential differences across demographic groups. Specifically, we investigated three key areas: (i) equity, by comparing whether prescription patterns varied by age, sex, or ethnicity; (ii) outcomes, by examining the relationship between patient outcomes, such as LOS or in-hospital death, and the types of medication prescribed; and (iii) public health insights, by analyzing whether drug types or dosages differed systematically across populations. We applied machine learning models to assess fairness across combinations of demographic characteristics, with the goal of supporting more equitable healthcare delivery and providing evidence to inform public health decision-making.

2. Literature review

2.1. Equity in drug prescriptions

Disparities in drug prescribing practices represent a significant violation of health equity, particularly given the well-documented impact of demographic variables (e.g., race, ethnicity, gender, and age) on the likelihood of receiving certain medications. These disparities not only reflect systemic injustice but also pose a long-term public health concern, contributing to unequal health outcomes. Torres[12] reported substantial racial differences in opioid prescribing. Specifically, historical healthcare data show that white patients were more likely to receive opioid medications for pain than Black or Hispanic patients, even when presenting with similar clinical symptoms. Idaho State University identified similar patterns, confirming this phenomenon despite increased awareness of these trends and their public health implications. In a meta-analysis by Wilson *et al*.,[13] racial and ethnic minorities were found to be routinely undertreated for pain, with significant differences observed across multiple clinical settings.

Gender-based prescribing differences have also been noted. Research shows that men typically receive higher doses of pain medications, whereas women are prescribed psychiatric drugs more frequently.[14] Implicit biases and limited time for clinical reflection may exacerbate these inequities in high-pressure environments such as the ICU. Age introduces an additional layer of complexity.

Evidence-based prescribing guidelines for older adults are limited, as this population is often underrepresented in clinical trials. As a result, prescribing practices, particularly for opioids and antibiotics, tend to be more conservative or arbitrary, potentially impacting health outcomes such as mortality and recovery time.[15] These verified disparities underscore the importance of integrating equity assessments into machine learning pipelines. Through advanced analytics, researchers can determine whether predictive algorithms may perpetuate existing inequities in health care. This is particularly relevant as clinical decision support systems are increasingly adopted in emergency departments and ICUs.

2.2. Pharmaceutical outcomes: Mortality and LOS

Patient outcomes, especially LOS and in-hospital mortality, are significantly influenced by the selection and timing of drug administration in critical care settings. For instance, antibiotics are often administered empirically before culture results are available to treat sepsis and other serious infections.[16] When appropriately prescribed, these medications can reduce mortality. However, overuse or delayed de-escalation may result in extended hospital stays, superinfections, and antibiotic resistance.

Opioids, while essential for pain management in critical care, pose their own risks. Long-term opioid usage has been linked to complications such as delirium, respiratory depression, and increased mortality in vulnerable populations.[17] In addition, patients discharged from ICUs may become dependent on opioids, exacerbating the ongoing opioid crisis. These risks highlight the need to examine both short- and long-term effects of drug administration.

Machine learning methods have shown considerable promise in identifying such patterns by analyzing large-scale EHR datasets—patterns that standard statistical models may overlook.[18] While more complex models such as random forests can enhance predictive accuracy, interpretable models such as logistic regression and decision trees are particularly valuable for understanding clinical relationships. The MIMIC-III and MIMIC-IV databases have been used in recent research to predict mortality and LOS using a variety of features, including laboratory results, medications, and vital signs.[19] However, few studies have clearly linked specific drug categories (such as opioids and antibiotics) to patient outcomes while also considering demographic differences. This gap in the literature is what the present study aims to address.

2.3. Public health insights and the role of EHRs

EHRs have become essential tools for public health monitoring, providing a high-fidelity, real-time view of clinical care, including prescription drug usage, diagnostic procedures, and outcomes. Among EHR databases, MIMIC-IV stands out for its comprehensiveness, openness, and depth. From a public health perspective, EHRs enable the identification of systemic prescribing patterns and potential inequities at the population level.[14,20-22] They can help answer critical questions such as: Are certain demographic groups more likely to receive particular medications? Are there gender- or ethnicity-based differences in dosage patterns? These insights are valuable not only for hospital management but also for policymakers and regulatory bodies responsible for ensuring equitable healthcare delivery.[21]

Furthermore, the incorporation of data visualization and machine learning tools has enhanced EHR-based research.[22] Fairness-aware learning frameworks allow researchers to measure and mitigate model bias across subgroups. For instance, equity metrics such as demographic parity or equal opportunity can be applied to identify and address performance gaps in mortality prediction models that disproportionately affect minority patients.[23] Despite these advances, translating model performance into actionable insights remains a significant challenge in public health analytics, especially when findings indicate systemic disparities in the provision of care. This underscores the need for ethical and policy-oriented frameworks to guide the interpretation and application of data-driven findings, in addition to technical rigor.[10,24]

Despite a growing body of work on pharmaceutical interventions and fairness in AI models, few studies have integrated these perspectives into a unified framework that links prescription patterns, outcomes, and equity in critical care settings.[25] The majority of the current literature tends to focus either on model performance or demographic disparities in isolation, without considering how predictive technologies may reinforce existing healthcare inequities.[26] Moreover, research that does address equity often relies on simplified prediction tasks or synthetic datasets, which lack the complexity of real-world clinical environments. By examining the relationship between drug type, demographic group, and patient outcomes using real ICU data from MIMIC-IV,[27,28] our work addresses this knowledge gap and provides a more meaningful understanding of how efficacy and equity intersect in critical care.

Disparities in medication prescriptions across demographic factors, including gender, race, and age, remain underexplored in large-scale EHR databases, even as data-driven approaches gain popularity in health care. Much of the current research on model fairness does not account for specific medications or focuses only on clinical outcomes such as mortality or LOS. However, the question

of whether medications, particularly high-impact drugs such as opioids and antibiotics, are prescribed fairly across different populations remains pressing. Furthermore, equity indicators are not often assessed in conjunction with downstream outcomes such as LOS and in-hospital mortality. This limits the translation of AI-driven findings into practical and ethically sound public health strategies. Various predictive models also neglect equity considerations, risking the reinforcement of existing biases in clinical care. If differences in model performance across demographic groups remain unexamined, such models could contribute to inequitable treatment decisions. The main contributions of our research are as follows:

(i) We conduct a detailed examination of prescription differences, especially for opioids and antibiotics, across demographic variables such as age, gender, and ethnicity, emphasizing the importance of equity in clinical decision-making

(ii) We investigate the associations between different types of prescriptions (e.g., opioids versus antibiotics) and key patient outcomes, including LOS and in-hospital mortality, to understand the broader impact of pharmaceutical choices

(iii) We identify prescription trends and demographic groups to determine whether certain populations receive different types or quantities of medications. These insights may inform evidence-based and more equitable public health policy.

3. Methodology

This research adopts a data-centric strategy to investigate fairness in treatment, medication effectiveness, and public health policy by applying advanced analytics to EHRs. We aim to evaluate the association between clinical interventions, specifically the use of drugs, and patient mortality using the publicly accessible MIMIC-IV dataset, while examining differences among demographic groups. Our methodology included data extraction, preprocessing, exploratory analysis, machine learning-based prediction modeling, and evaluation. The basic analytical pipeline was built using Python and tools such as pandas, scikit-learn, statsmodels, seaborn, and matplotlib. To forecast patient mortality, models such as decision tree classifiers were used. Figure 1 shows the proposed architecture.

3.1. Dataset description

The MIMIC-IV[27,28] is a comprehensive, openly available database containing de-identified health data from more than 70,000 ICU patients at Beth Israel Deaconess Medical Centre from 2008 to 2019. This extensive database includes a wide range of clinical data, such as patient demographics, diagnoses, medications, procedures, vital signs, test results, and mortality outcomes. For this study, a subset of the MIMIC-IV database was meticulously retrieved and vetted to focus solely on adult ICU patients. Demographic characteristics, specifically age, gender, and race, were

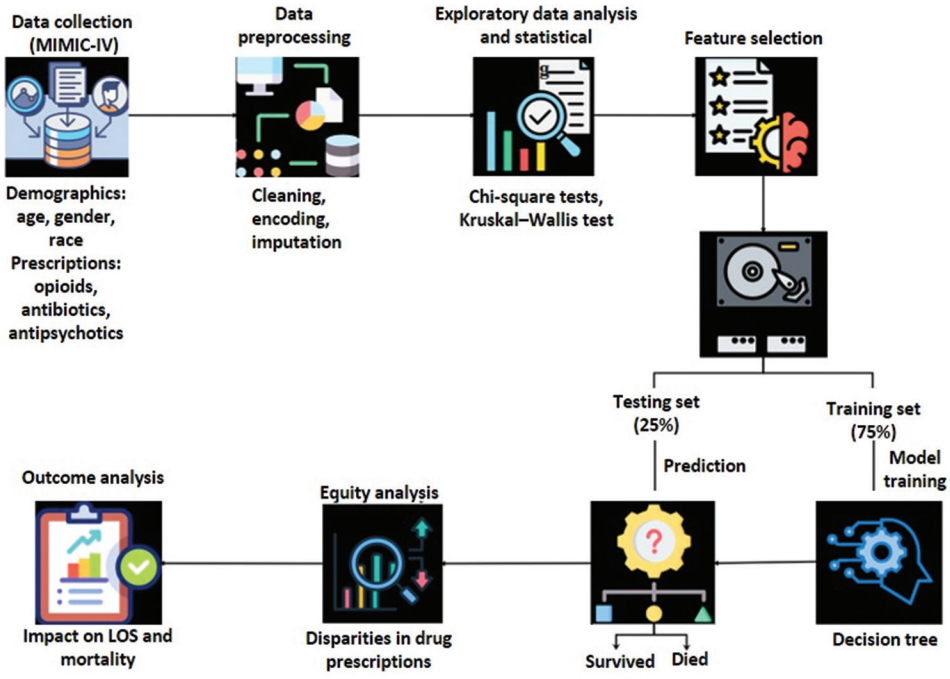

Figure 1. Proposed architecture
Abbreviations: LOS: Length of stay; MIMIC-IV: Medical Information Mart for Intensive Care IV.

among the key factors chosen for analysis. The main outcome of interest is in-hospital mortality, represented as a binary indicator—0 for individuals who survived and 1 for those who passed away during hospitalization. This carefully assembled dataset serves as a basis for examining the relationship between medical treatments and patient outcomes, whereas also allowing for the investigation of differences across different demographic groups.

3.2. Data preprocessing

Before conducting any modeling or analysis, the dataset underwent several preprocessing steps to ensure it was clean, consistent, and suitable for statistical and machine learning applications. Given that most algorithms require numerical input, gender and race were encoded from categorical to numeric format. As part of data preprocessing, the age variable was grouped into categories to facilitate demographic analysis and assess fairness. The continuous age values were divided into three bins: "young adult" (18–35 years), "adult" (36–65 years), and "senior" (66 years and older). This grouping simplified comparisons between different age brackets in subsequent analyses and made patterns related to age easier to interpret. To avoid issues during model training, the binary treatment indicator was also converted into an integer format. To preserve data integrity, records lacking important information such as demographics, treatment status, or outcome data were either eliminated or appropriately imputed, depending on the context. A carefully selected set of pertinent features was chosen to provide a strong basis for analysis while balancing simplicity and predictive power. Finally, the dataset was divided into training and testing sets in a 75:25 ratio to ensure an unbiased and equitable assessment of the model's performance.

As part of preprocessing, the age variable was grouped into three bins (18–35, 36–65, and 66+). While this categorization simplified comparisons, it may obscure finer age-related prescribing patterns, a limitation we acknowledge later. Records with >20% missing critical information (e.g., demographics, treatment status, or outcomes) were removed, whereas those with <20% missing values were imputed using median values for continuous variables and mode values for categorical variables. Gender and race were encoded using one-hot encoding, and binary treatment indicators were recoded as 0/1 integers. Outlier checks were performed using the interquartile range method for age, LOS, and laboratory measures, and implausible extreme values were excluded. Encoded variables were validated against raw data to ensure accuracy and consistency.

3.3. Exploratory data analysis

Exploratory data analysis was used to gain better insight into the structure of the dataset, as well as the relationships among important variables. One of the initial analyses involved examining the target variable and understanding its distribution to establish the proportion of individuals who survived versus those who did not. While there was a slight class imbalance, the results supported the decision to implement robust machine learning models to address this imbalance in future analyses. The visualization also provided insights into whether there was a noticeable correlation between the length of time in treatment and mortality. The study further examined the relationships between demographic factors such as gender and race, and patient outcomes. Using statistical and visual methods to investigate differences between demographic groups provided preliminary ideas for subsequent steps focused on fairness analysis. Finally, the study conducted a correlation analysis between features and target variables, which serves two purposes: first, to identify potential multicollinearity, and second, to determine the most influential determinants of the outcome variable in the dataset.

3.4. Predictive modeling

Supervised learning techniques were employed to explore the association between prescription drug use, patient demographics, and in-hospital mortality. The decision tree classifier was selected for its interpretability and its ability to its non-linear relationships. Its capacity to handle both numerical and categorical variables makes it a suitable choice for clinical datasets, which often contain complex relationships and noise. The model's primary features included variables related to medication use and demographic factors such as age, gender, and race.

The dependent variable represented a binary outcome indicating the patient's survival status. Stratification was applied during the train–test split to preserve the class distribution of the survival outcome across both subsets. The dataset was split into training and testing sets using a 75:25 ratio. The decision tree model was trained on selected features using default hyperparameters and then evaluated on the test set. Consistent results were observed when the model was trained using a random seed for performance diagnostics. To assess the model's classification performance, the metrics below were calculated:

(i) Accuracy: Percentage of instances correctly categorized.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{(I)}$$

(ii) Precision: Percentage of true positive (TP) predictions out of all positive predictions.

$$Precision = \frac{TP}{TP + FP} \quad \text{(II)}$$

(iii) Recall (sensitivity): Percentage of TP identified among all actual positives.

$$Recall = \frac{TP}{TP + FN} \quad \text{(III)}$$

(iv) F1-score: Harmonic mean of precision and recall, providing a balance between the two.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \text{(IV)}$$

False positives (FP) are incorrectly identified positive cases, true negatives (TN) are correctly identified negative cases, TP are correctly identified positive cases, and false negatives (FN) are positive cases that the model failed to identify. Together, these metrics provide a comprehensive view of the model's classification performance, accuracy, and reliability.

4. Results and discussion

As healthcare datasets grow in scale and diversity, protecting patient privacy while ensuring analytic utility becomes critical. Recent research on robust steganography techniques for medical records offers innovative ways to safeguard sensitive patient data during healthcare analytics.[29] Beyond security, emerging self-supervised learning frameworks demonstrate the ability to capture complex, high-dimensional patterns in clinical data while potentially maintaining interpretability. Incorporating these advanced approaches may enhance future work in equity-focused health informatics, enabling richer insights without compromising ethical or privacy considerations.[30]

4.1. Equity: Disparities in drug prescriptions

We investigated the distribution of prescription drug types, specifically opioids, antibiotics, and antipsychotics, across key demographic factors such as gender, race, and age group, to evaluate equity in prescribing practices. This analysis aimed to identify potential biases or inequities in medical decision-making. Our analysis showed that opioids were prescribed more frequently to men than to women, whereas women received antibiotics at a slightly higher rate. Antipsychotic prescriptions were nearly equally distributed between genders. These results suggest that drug selection may be influenced by latent clinical biases or gender-based preferences. The statistical significance of the association between drug type and gender was confirmed by a chi-square test ($\chi^2 = 9.39$, p=0.00916), indicating that gender plays a role in prescription patterns. Figure 2 shows the percentage distribution of prescribed drug types by gender.

Patients' races were categorized into four groups: White, Black, Hispanic, and Unknown. White patients received more opioid prescriptions than other racial groups, while Black and Hispanic patients were prescribed antibiotics at comparatively higher rates. Black patients also had a slightly higher percentage of antipsychotic use, although antipsychotic prescriptions were more uniformly distributed overall. Race had a substantial impact on prescribing trends, as confirmed by the chi-square test ($\chi^2 = 75.60$, p<0.00001). Figure 3 illustrates the percentage distribution of prescribed drug types by race.

Different prescribing patterns were observed across age groups. Opioids were more frequently administered to young individuals (18–35 years old), which may reflect differences in clinical presentation or prescribing preferences. Antibiotics were disproportionately prescribed to older adults (65 years and older), likely due to a higher risk of infection. The majority of antipsychotic prescriptions were given to middle-aged individuals (35–65 years old). These age-specific prescribing trends were statistically significant, as confirmed by the Chi-square test (χ^2 = 31.24, p<0.00001), suggesting the potential influence of unconscious bias or age-related clinical recommendations. Figure 4 depicts the percentage distribution of prescribed drug types by age group.

The percentage distribution of prescriptions for three drug classes—opioids, antibiotics, and antipsychotics—across four racial groups (Black, Hispanic, Unknown, and White) is shown in Figure 5. The chart highlights significant racial variance

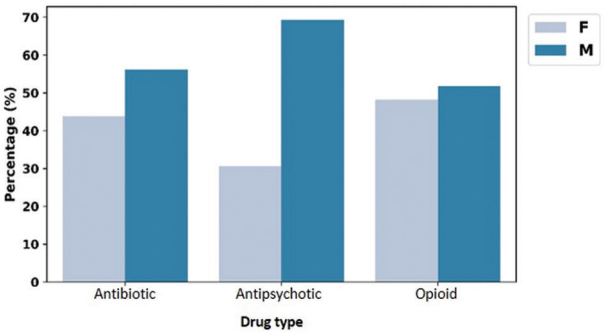

Figure 2. Percentage distribution of prescribed drug types by gender
Abbreviations: F: Female; M: Male.

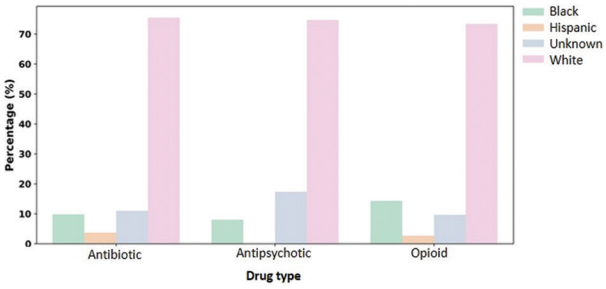

Figure 3. Percentage distribution of prescribed drug types by race

in prescription trends, which may reflect underlying inequities in healthcare delivery and clinical decision-making processes.

Opioids account for the highest percentage of prescriptions across all racial groups, with Black patients having the highest percentage, closely followed by White and unknown individuals. This finding contradicts previous research suggesting that Black patients were underprescribed opioids, indicating either a shift in clinical practice or dataset-specific factors. More than half of all prescriptions for Hispanic patients were antibiotics, which may point to a higher rate of infection-related illnesses or possible issues of overprescription. Interestingly, White patients and those in the unknown race category had a more balanced distribution of opioid and antibiotic prescriptions, with a relatively small percentage of antipsychotic usage (~8–10%) across all groups, except for Hispanic patients, where antipsychotic use appeared lowest. Such disparities imply that prescription behavior may systematically vary by race, possibly due to differences in diagnostic patterns, provider bias, or dynamics related to access to care. Further research is necessary to determine whether the high rates of opioid prescriptions among Black patients and the elevated rates of antibiotic prescriptions among Hispanic patients reflect genuine clinical needs or are indicative of systemic injustices.

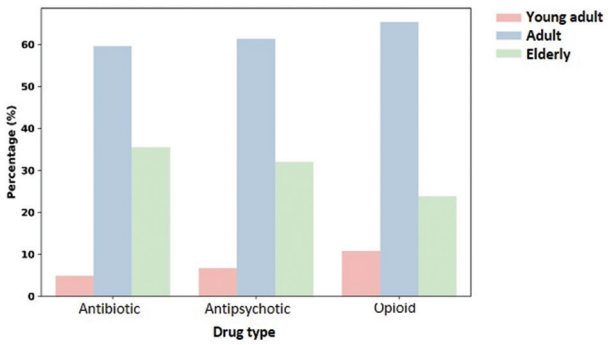

Figure 4. Percentage distribution of prescribed drug types by age group

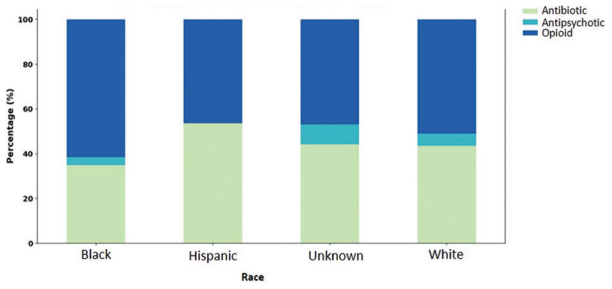

Figure 5. Racial differences in proportions of prescribed drug categories

4.2. Pharmaceutical prescription outcomes: Impact of drug type on LOS and in-hospital mortality

To evaluate the effect of pharmaceutical prescription type on patient LOS, we classified patients according to the primary medication they received, focusing on opioids, antibiotics, and antipsychotics. Due to the non-normal distribution of LOS data, the Kruskal–Wallis H-test, a non-parametric alternative to Analysis of Variance, was used. The results showed a statistically significant difference in LOS across drug types (H = 14.78, p=0.0006), suggesting that the type of medication influences the duration of hospitalization.

However, no *post hoc* test (e.g., Dunn's test with Bonferroni correction) was performed, limiting the interpretation of specific group differences. Moreover, adjustments for multiple comparisons were not made, increasing the risk of Type I error. Furthermore, key confounding variables such as comorbidities, illness severity, and socioeconomic status could not be controlled due to limitations of the dataset.

Antibiotic prescriptions were associated with the shortest hospital stays, while antipsychotic prescriptions corresponded to the longest average LOS, followed by opioid prescriptions, as shown in Figure 6. These findings highlight the clinical and operational significance of pharmacological choices, not only in terms of treatment effectiveness but also regarding healthcare costs and resource usage. Shorter LOS in patients treated with antibiotics may reflect acute, rapidly resolved illnesses, whereas longer LOS among those prescribed antipsychotics may point to the complex and chronic nature of psychiatric conditions.

We also examined the relationship between primary drug type and in-hospital mortality, a key clinical outcome, using the chi-square test of independence. Descriptive analysis indicated that patients prescribed antibiotics had the highest mortality rate, followed by those receiving antipsychotics. The group prescribed opioids had the lowest mortality rate.

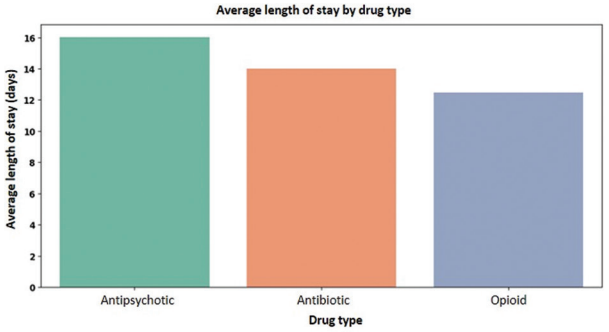

Figure 6. Impact of drug type on length of stay

The chi-square test produced a non-significant result ($\chi^2 = 1.69$, p=0.4291), indicating no statistically significant association between prescription type and in-hospital mortality. Although visual trends pointed to possible variations, these were not supported by inferential statistics and may be influenced by unmeasured patient characteristics or clinical contexts not adequately captured by drug category alone.

Although mortality rates varied slightly by drug type, they did not differ significantly at the population level, as illustrated in Figure 7. These findings underscore the importance of future multivariable modeling that incorporates comorbidities, illness severity, and treatment protocols, as drug type alone may not be a reliable predictor of mortality risk.

To further investigate potential differences in in-hospital mortality across pharmacological categories, we analyzed the percentage of deaths and survivals among patients treated with opioids, antipsychotics, and antibiotics. The distribution of mortality status (0 = survived, 1 = died) for each drug type group is shown in Figure 8.

The analysis revealed relatively similar survival rates across all drug categories. In particular, patients prescribed opioids had the lowest mortality rate (~15%) and the highest survival rate (~85%). On the other hand, the antipsychotic and antibiotic groups showed lower survival rates (around 83%) and slightly higher mortality rates (approximately 17%). These results suggest that while mortality rates differ slightly among drug categories, the differences are not substantial. The lower mortality rate observed in the opioid group may be due to factors such as variation in clinical indication, illness severity, or patient comorbidities, rather than the direct effect of the medication itself.

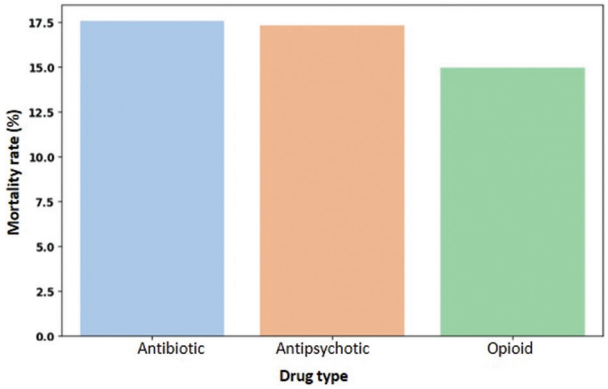

Figure 7. Impact of drug type on mortality rates

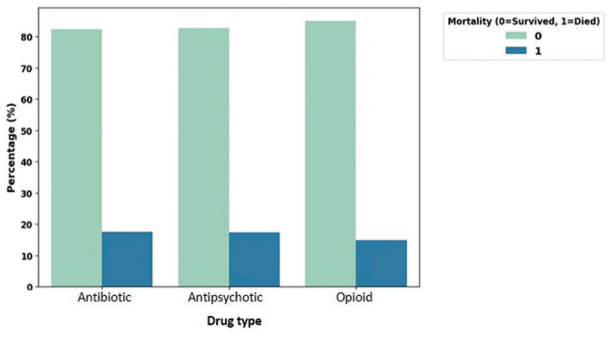

Figure 8. Mortality distribution by drug type

4.3. Public health decision-making: Predicting in-hospital mortality using a decision tree classifier

The decision tree classifier demonstrated strong performance across several evaluation metrics after being trained and assessed on the MIMIC-IV clinical dataset. Due to the imbalance in the survival outcome classes, the F1-score was prioritized to ensure that both precision and recall were adequately considered during the evaluation. Table 1 presents a comparison of machine learning models for in-hospital mortality prediction.

In identifying in-hospital mortality cases specifically, the decision tree's F1-score of 0.8235 represented a balanced trade-off between recall and precision. Moreover, its overall accuracy—93.47% in classifying both survival and non-survival instances—highlighted the model's decisional capability. The model's notable discriminative ability was further confirmed by an area under the receiver

Table 1. Performance comparison of machine learning models for in-hospital mortality prediction

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
CatBoost	92.39	0.8163	0.7692	0.7921	0.9722
LightGBM	91.67	0.7458	0.8462	0.7928	0.9694
XGBoost	91.67	0.7458	0.8462	0.7928	0.9685
Random forest	92.03	0.8000	0.7692	0.7843	0.9615
Gradient boosting	91.67	0.7959	0.7500	0.7723	0.9490
K-nearest neighbors	89.49	0.6885	0.8077	0.7434	0.9239
AdaBoost	84.06	0.5870	0.5192	0.5510	0.8734
Logistic regression	80.07	0.4211	0.1538	0.2254	0.8639
SVM	81.16	0.0000	0.0000	0.0000	0.8758
Decision tree	93.48	0.8400	0.8077	0.8235	0.9337

Abbreviations: AdaBoost: Adaptive boosting; AUC: Area under the curve; CatBoost: Categorical Boosting; LightGBM: Light gradient boosting machine; ROC: Receiver operating characteristic; SVM: Support vector machine; XGBoost: eXtreme Gradient Boosting.

operating characteristic curve (ROC-AUC) score of 0.9337, indicating a strong capacity to differentiate between patients who survived and those who did not.

To evaluate the robustness of the proposed model, results were compared with several machine learning algorithms: random forest, gradient boosting, adaptive boosting (AdaBoost), eXtreme Gradient Boosting (XGBoost), light gradient boosting machine (LightGBM), k-nearest neighbors (KNN), support vector machine (SVM), logistic regression, and categorical Boosting (CatBoost). The ensemble methods—CatBoost, LightGBM, and XGBoost—yielded ROC-AUC scores above 0.96 and F1-scores of 0.7921, 0.7928, and 0.7927, respectively. Although these ensemble models outperformed all others, they are often more complex and less interpretable.

Conversely, the decision tree exhibited a modestly superior F1-score while preserving a high degree of interpretability. This is particularly important in healthcare settings, where model transparency and explainability are critical. The simple design also fosters greater trust and facilitates practical implementation, as the tree structure allows physicians to follow the decision pathway and understand the contributing factors behind each prediction. In contrast, linear models such as SVM and logistic regression perform poorly. The SVM failed to detect any positive cases (F1-score = 0.0), while logistic regression achieved a relatively low F1-score of 0.225. These results suggest that the complex, non-linear relationships present in the demographic and medication data cannot be captured effectively using linear decision boundaries.

Overall, the evaluation findings indicate that the decision tree classifier is a satisfactory and efficient model for predicting in-hospital mortality. It offers a reasonable option for supporting clinical decisions, with promising predictive accuracy and interpretability, while maintaining performance reliability, as depicted in Figure 9.

Figure 10 shows the confusion matrix and ROC curve for the decision tree, used to assess the model's performance.

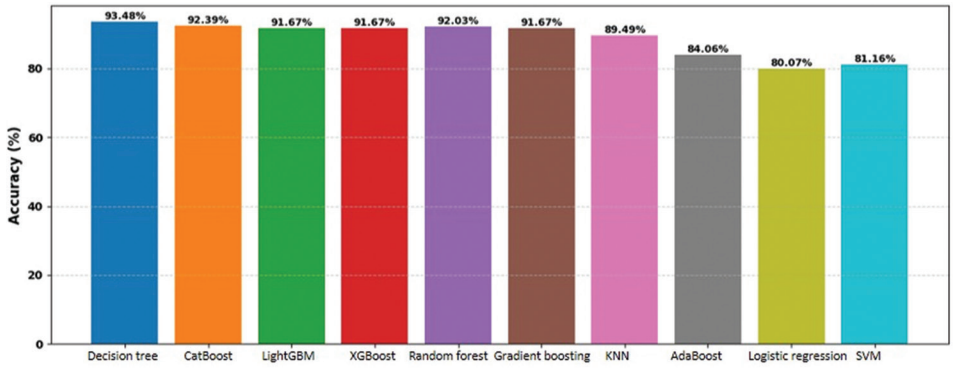

Figure 9. Accuracy comparison of the proposed model with comparative approaches
Abbreviations: AdaBoost: Adaptive boosting; CatBoost: Categorical Boosting; KNN: K-nearest neighbors; LightGBM: Light Gradient Boosting Machine; SVM: Support vector machine; XGBoost: eXtreme Gradient Boosting.

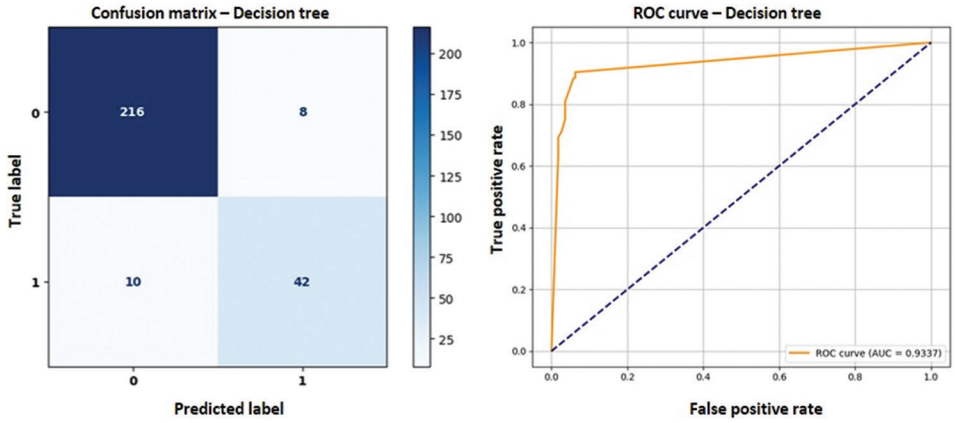

Figure 10. Confusion matrix and receiver operating characteristic (ROC) curve for the proposed decision tree
Abbreviation: AUC: Area under the curve.

The confusion matrix reveals that the model misclassified 10 occurrences of class 1 as class 0 (FN) and eight instances of class 0 as class 1 (FP), while correctly classifying 216 instances of class 0 and 42 instances of class 1. These results yield an overall accuracy of approximately 93.48%, with a precision of about 84.00% and a recall of around 80.77% for class 1. The F1-score, which balances precision and recall, is approximately 82.35%, indicating that the model maintains a strong trade-off between sensitivity and specificity. Furthermore, the ROC curve illustrates the model's ability to distinguish between the two classes across various thresholds. The AUC value of 0.9337 reflects a very high level of discriminative performance. The model achieves a high TP rate while maintaining a low FP rate, suggesting that the decision tree classifier has strong predictive performance on this dataset.

Beyond performance metrics, we also examined model interpretability. The most important predictors were age, opioid prescription, and race. For example, patients >65 years with antipsychotic prescriptions had the highest mortality risk, while patients <35 years with antibiotic prescriptions had the lowest risk. Although ensemble methods such as CatBoost and LightGBM achieved higher AUC scores, their interpretability was limited. This study used a single 75:25 train–test split; future work should incorporate k-fold cross-validation and evaluate model calibration (e.g., Brier score, calibration plots).

5. Limitations and future work

While this study provides valuable insight into differences in prescription drug use and their relationship to patient outcomes, several important limitations should be noted. First, although the MIMIC-IV dataset is relatively large and comprehensive, the data used in this analysis were limited to a single hospital system and may not be generalizable to broader or more heterogeneous populations. Second, the assessment of medication exposure may be subject to measurement error, as the prescription data from MIMIC-IV reflect the documented medication orders, but not patient adherence or actual administration.

Additional methodological limitations should also be acknowledged. The broad age categorization into three groups simplified the analysis but may have obscured finer age-related prescribing patterns. Missing data were imputed using median and mode values, which may not accurately reflect the underlying distributions. Our statistical testing did not include *post hoc* procedures for the Kruskal–Wallis test and did not adjust for multiple comparisons, increasing the risk of FP. Moreover, confounding variables such as comorbidities, disease severity, and socioeconomic status were not available in MIMIC-IV, limiting our ability to control for key sources of bias. Predictive modeling relied on a single train–test split without cross-validation, and no model calibration was performed, which constrains the clinical applicability of the results.

Future studies should address these limitations by incorporating multicenter datasets to enhance generalizability and robustness. Expanding the analysis to include more detailed clinical variables and temporal information (e.g., medication administration timing, disease progression) may improve both the predictive and descriptive components of the research. In addition, applying fairness-aware machine learning techniques could help detect and mitigate potential biases in algorithmic predictions, ultimately supporting the development of more equitable and robust clinical decision support systems.

6. Conclusion

This study utilized the MIMIC-IV dataset to thoroughly investigate differences in pharmaceutical prescribing characteristics and how those characteristics impact clinical outcomes. We examined three broad topics: the use of machine learning for predicting mortality risk, the impact of drug type on LOS and in-hospital mortality, and equity in the distribution of prescription drugs based on demographics. Our research found statistically significant differences in prescribing behavior by age, race, and gender. While antibiotics were prescribed more frequently to women, pain medications were more likely to be prescribed to men. Racial differences were more pronounced, with African American patients receiving more pain medication than previously reported, and Hispanic patients receiving more antibiotics. These trends, which are supported by statistically significant findings, highlight the potential for bias or systemic influences in clinical judgment and underscore the pressing need for equitable healthcare delivery.

In addition, we found that the type of medication prescribed has a substantial impact on LOS—antipsychotics were associated with the longest hospital stays, whereas antibiotics were linked to shorter stays. Nevertheless, no statistically significant correlation was found between drug type and in-hospital mortality, indicating that factors beyond prescription type likely contribute to mortality outcomes. In terms of predictive modeling, the decision tree classifier proved to be a robust and interpretable tool for predicting in-hospital mortality, achieving an F1-score of 0.8235 and an ROC-AUC of 0.9337. While ensemble techniques such as CatBoost and LightGBM slightly outperformed it in terms of AUC, the transparency and explainability of the decision tree make it particularly

suitable for clinical applications, where understanding the rationale behind predictions is crucial. The results of this study highlight the importance of addressing inequities in prescription practices and demonstrate how explanatory machine learning models can support clinical decision-making. Future research should aim to incorporate a broader range of clinical variables (e.g., comorbidities, treatment protocols) to better identify the drivers of outcomes and to design equitable, data-driven interventions that optimize both the effectiveness and equality of healthcare systems.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Author contributions

Conceptualization: Md. Majedur Rahman, Md. Shihab Rahman, Ahmed Tanvir

Data curation: Safiul Islam, Sajidul Islam Khan

Formal analysis: Abdullah Al Mahmud Ashik, Emran Hossain

Methodology: Ahmed Tanvir, Md. Majedur Rahman, Safiul Islam

Writing–original draft: Md. Majedur Rahman, Md. Shihab Rahman, Safiul Islam, Sajidul Islam Khan, Abdullah Al Mahmud Ashik, Emran Hossain

Writing–review & editing: Md. Majedur Rahman, Ahmed Tanvir

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

All datasets generated during the present study are included in the article and/or the supplementary materials. Further inquiries can be directed to the corresponding authors.

References

1. Muthukaruppasamy S, Thomas AS, Gopal JN, Anand AJ, Parimalasundar E. Diagnostics, treatment and patient care in the age of digital twins: A game-changer in healthcare. In: *Handbook of Industrial and Business Applications with Digital Twins*. United States: CRC Press; 2025. p. 282-305.
2. Bhuiyan MNA, Mondal RS. AI-driven predictive analytics in healthcare: Evaluating impact on cost and efficiency. *J Comput Anal Appl*. 2023;31(4):1355-1371.

   doi: 10.48047/jocaaa.2023.31.04.26
3. Abbas SR, Abbas Z, Zahir A, Lee SW. Federated learning in smart healthcare: A comprehensive review on privacy, security, and predictive analytics with IoT integration. *Healthcare (Basel)*. 2024;12(24):2587.

   doi: 10.3390/healthcare12242587
4. Marcos-Vidal JM, González R, Merino M, Higuera E, García C. Sedation for patients with sepsis: Towards a personalised approach. *J Pers Med*. 2023;13(12):1641.

   doi: 10.3390/jpm13121641
5. Green S, Prainsack B, Sabatello M. Precision medicine and the problem of structural injustice. *Med Health Care Philos*. 2023;26(3):433-450.

   doi: 10.1007/s11019-023-10173-9
6. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. *MIMIC-IV. PhysioNet*. 2020. p. 49-55.

   doi: 10.13026/s6n6-xd98
7. Hanna MG, Pantanowitz L, Jackson B, *et al*. Ethical and bias considerations in artificial intelligence/machine learning. *Mod Pathol*. 2025;38(3):100686.

   doi: 10.1016/j.modpat.2024.100686
8. Bhuiyan MNA, Kamruzzaman M, Saha S, Siddiki MS, Mondal RS. Role of data analysis and integration of artificial intelligence. *J Bus Manag Stud*. 2025;7(4):379-388.

   doi: 10.32996/jbms.2025.7.4.20.26
9. Zaidi SAJ, Ghafoor A, Kim J, Abbas Z, Lee SW. HeartEnsembleNet: An innovative hybrid ensemble learning approach for cardiovascular risk prediction. *Healthcare (Basel)*. 2025;13(5):507.

   doi: 10.3390/healthcare13050507
10. Kamruzzaman M, Mondal RS, Islam MK, Rahaman MA, Saha S. AI-driven predictive modelling of US economic growth using big data and explainable machine learning. *Int J Comput Exp Sci Eng*. 2024;10(4):1927-1938.

    doi: 10.22399/ijcesen.3612.
11. Mondal RS, Kamruzzaman M, Saha S, Bhuiyan MNA. Quantum machine learning approaches for high-dimensional cancer genomics data analysis. *Comput Integr Manuf Syst*. 2025;31(1):13-32.

    doi: 10.24297/j.cims.2025.1.21
12. Torres PJ. Race and gender disparities in pain treatment and opioid prescribing. *Soc Sci Med*. 2025;374:118011.

    doi: 10.1016/j.socscimed.2025.118011
13. Wilson M, Booker S, Saravanan A, *et al*. Disparities,

inequities, and injustices in populations with pain: Nursing recommendations supporting ASPMN's 2024 position statement. *Pain Manag Nurs*. 2025;26(2):139-148.

doi: 10.1016/j.pmn.2024.10.016

14. Yang N, Fang MC, Rambachan A. Sex disparities in opioid prescription and administration on a hospital medicine service. *J Gen Intern Med*. 2024;39:2679-2688.

doi: 10.1007/s11606-024-08814-7

15. Khan NF, Bykov K, Glynn RJ, Vine SM, Gagne JJ. Comparative risk of opioid overdose in patients who initiated antibiotics for urinary tract infection while on long-term opioid therapy. *Am J Epidemiol*. 2025;194(3):674-679.

doi: 10.1093/aje/kwae248

16. Qian J, Yuan Y, Shang Z, *et al*. Association between emergency department to intensive care units time and in-hospital mortality: An analysis of the MIMIC-IV database. *BMJ Open*. 2025;15(4):e090011.

doi: 10.1136/bmjopen-2024-090011

17. Ekong M, Monga TS, Daher JC, *et al*. From the intensive care unit to recovery: Managing post-intensive care syndrome in critically ill patients. *Cureus*. 2024;16(5):e61443.

doi: 10.7759/cureus.61443

18. Li C, Mowery DL, Ma X, *et al*. Realizing the potential of social determinants data in EHR systems: A scoping review of approaches for screening, linkage, extraction, analysis, and interventions. *J Clin Transl Sci*. 2024;8(1):e147.

doi: 10.1017/cts.2024.571

19. Gabitashvili A, Kellmeyer P. *Predicting Length of Stay in Neurological ICU Patients using Classical Machine Learning and Neural Network Models: A Benchmark Study on MIMIC-IV* [arXiv Preprint]; 2025.

doi: 10.48550/arXiv.2505.17929

20. Silva JM, Durden TE, Hirsch A. Erasing inequality: Examining discrepancies between electronic health records and patient narratives to uncover perceived stigma and dismissal in clinical encounters. *Soc Sci Med*. 2023;323:115837.

doi: 10.1016/j.socscimed.2023.115837

21. Williamson SM, Prybutok V. Balancing privacy and progress: A review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare. *Appl Sci*. 2024;14(2):675.

doi: 10.3390/app14020675

22. Ramakrishnaiah Y, Macesic N, Webb GI, Peleg AY, Tyagi S. EHR-ML: A data-driven framework for designing machine learning applications with electronic health records. *Int J Med Inform*. 2025;196:105816.

doi: 10.1016/j.ijmedinf.2025.105816

23. Percival V, Thoms OT, Oppenheim B, *et al*. The lancet commission on peaceful societies through health equity and gender equality. *Lancet*. 2023;402(10413):1661-1722.

doi: 10.1016/S0140-6736(23)01348-X

24. Kamruzzaman M, Saha S, Siddiki MS, Mondal RS, Bhuiyan MNA. Applications of artificial intelligence in small and medium scale business. *J Bus Manag Stud*. 2025;7(4):314-325.

doi: 10.32996/jbms.2025.7.4.20.21

25. Ghazalbash S, Zargoush M, Verter V, Perri D. Data-driven insights into interhospital care fragmentation: Implications for health policy and equity among older adults. *PLoS One*. 2025;20(2):e0316829.

doi: 10.1371/journal.pone.0316829

26. Chung RYN, Freedman B. Health inequalities in AI machine learning. In: *Artificial Intelligence in Medicine*. United States: Academic Press; 2024. p. 119-130.

27. Johnson A, Bulgarelli L, Pollard T, *et al*. *MIMIC-IV PhysioNet. RRID: SCR_007345*. Ver. 3.1. [Preprint]; 2024.

28. Johnson AEW, Bulgarelli L, Shen L, *et al*. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10:1.

doi: 10.1038/s41597-022-01899-x

29. Saha S, Islam MK, Rahaman MA, Mondal RS, Kamruzzaman M. Machine learning driven analytics for national security operations: A wavelet-stochastic signal detection framework. *J Comput Anal Appl*. 2024;33(8):210.

doi: 10.48047/jocaaa.2024.33.08.210

30. Majchrowska S, Hildeman A, Mokhtari R, Diethe T, Teare P. Exploring interpretable echo analysis using self-supervised parcels. *Comput Biol Med*. 2025;192:110322.

doi: 10.1016/j.compbiomed.2025.110322