Project Report

Effect of Early Vasopressor Administration on Septic Shock Mortality

Subject: STAT 432

Team Members:
Alarsh Tiwari (NetID: alarsht2)
Anshika Pradhan (NetID: anshika6)

Project Goal

This project aims to assess the impact of administering vasopressors, specifically norepinephrine, within one hour of septic shock onset on patient mortality rates. The primary objective is to provide evidence-based insights for optimizing ICU treatment strategies for septic shock patients.

Approach

The analysis utilized the MIMIC-IV dataset to identify a cohort of septic shock patients meeting specific clinical criteria. Patients were categorized into treatment and control groups based on the timing of vasopressor administration. Propensity score matching (PSM) was applied to balance baseline covariates between groups, ensuring fair comparisons. Multiple causal inference techniques, including Inverse Probability Weighting (IPW) and DoWhy analysis, were used to estimate the causal effect of early vasopressor administration.

Unique Methods

1. **Direct Covariate Matching**: Beyond propensity score matching, this project employed direct covariate matching to enhance the robustness of the comparison.
2. **Advanced Causal Inference Tools**: Python-based tools like the DoWhy package were utilized for validating the causal relationships, providing deeper insights into treatment effects.
3. **Extended Feature Engineering**: Feature selection was driven by domain relevance and statistical importance, using a random forest classifier to identify predictors with $>1.2\%$ importance.
4. **Custom Imputation Techniques**: Missing lab event values were imputed with KNN, while categorical variables were imputed with the mode to ensure comprehensive data integrity.

AI Usage Statement

Disclaimer

All AI tools and methodologies were utilized in accordance with the university's academic integrity policy. These tools served as supplementary aids, with all critical decisions, interpretations, and implementations carried out by the project team.

In completing this project, the following AI tools and techniques were utilized:

Code Assistance: AI was used to debug and optimize Python code, especially in the implementation of the Propensity Score Matching (PSM) and causal inference methods.

Literature Insights: AI assisted in summarizing relevant literature, highlighting gaps that this project aimed to address.

Language Refinement: AI-based grammar tools assisted in refining the report text to ensure clarity, coherence, and academic tone.

Conversion to latex formatting: Chatgpt helped in the conversion of raw text to latex, wherever needed.

Literature Review

1. Timing of Vasopressor Initiation and Mortality in Septic Shock

Authors: Hernández G, Ospina-Tascón GA, Damiani LP, et al.
Published In: Critical Care, 2020
Dataset: Multicenter cohort study with 6,514 septic shock patients. Data included timing of vasopressor administration, APACHE II severity scores, and baseline comorbidities.
Findings:

- **Mortality Trends**: Mortality increased with delays in vasopressor initiation, with the first hour being critical for improved survival.
- **Temporal Analysis**: A decile-based stratification revealed a consistent rise in mortality from 47.6% in the earliest decile to 63% in the most delayed group.
- **Key Predictors**: Higher APACHE II scores and antimicrobial delays were strongly associated with mortality, highlighting the need for timely interventions beyond vasopressors.

Relevance to our Project: This study provides a framework for analyzing the timing effects of norepinephrine. However, it lacks causal inference models like DoWhy, creating opportunities for our project to offer additional insights.
URL: https://doi.org/10.1186/s13613-020-00687-0

2. Very Early Start of Norepinephrine in Septic Shock

Authors: Ospina-Tascón GA, Hernández G, Alvarez I, et al.
Published In: Critical Care, 2020
Dataset: Latin American ICU registry with detailed records on norepinephrine administration and patient outcomes.
Findings:

- **Propensity Score Matching (PSM)**: Matched patients receiving norepinephrine within the first hour against those who received it later. Early administration led to a 15% reduction in 28-day mortality.
- **Outcome Predictors**: Serum lactate levels >2 mmol/L and MAP <65 mmHg at diagnosis were critical in identifying patients needing immediate vasopressor support.
- **Sensitivity Analysis**: Results remained consistent even after using gradient-boosted models for alternative PSM.

Relevance to our Project: Aligns with our use of PSM but does not employ advanced causal techniques or examine alternative time thresholds beyond one hour. **URL**: https://doi.org/10.1186/s13613-020-00688-z

3. Early Versus Delayed Administration of Norepinephrine in Septic Shock

Authors: Bai X, Yu W, Ji W, et al.
Published In: Critical Care, 2014
Dataset: Single-center hospital registry with data on norepinephrine administration timing and patient outcomes.
Findings:

- **Survival Analysis**: Early norepinephrine (within 3 hours) reduced 28-day mortality by 20%.
- **Predictive Markers**: Elevated lactate and critically low MAP were strong predictors of poor outcomes without timely intervention.
- **Exploratory Observations**: Patients receiving norepinephrine earlier also required less total fluid resuscitation, reducing the risk of fluid overload.

Relevance to our Project: Complements your work by reinforcing the importance of early norepinephrine but lacks causal techniques like IPW or PSM.
URL: https://doi.org/10.1186/s13613-014-0031-0

4. Timing of Vasopressin Initiation and Mortality in Septic Shock

Authors: BMC Infectious Diseases, 2023
Dataset: MIMIC-III v1.4 and MIMIC-IV v2.0 databases, analyzing 1,817 septic shock patients.
Findings: - **Doubly Robust Analysis**: Combined PSM with logistic regression to estimate causal effects. Patients initiating vasopressin with low-dose norepinephrine (<0.25 µg/kg/min) had a 34% lower 28-day mortality compared to those at high doses.

- **Key Observations**: Early vasopressin initiation reduced fluid overload and lengthened CRRT-free and ventilation-free days.
- **Statistical Techniques**: Used advanced sensitivity analyses to validate findings across multiple causal models.

Relevance to our Project: Directly aligns with your dataset (MIMIC-IV) but focuses on vasopressin rather than norepinephrine timing, allowing for a complementary rather than overlapping analysis.
URL: https://doi.org/10.1186/s12879-023-08561-5

Dataset

The MIMIC-IV (Medical Information Mart for Intensive Care IV) dataset is a vast, publicly available resource that contains detailed clinical data for patients admitted to critical care units. It spans multiple years of data from Beth Israel Deaconess Medical Center and is used extensively in healthcare research, particularly for studying intensive care units (ICU) and patient outcomes. This dataset provides a rich set of variables that reflect the complex nature of ICU treatment and the critical care environment. MIMIC-IV is organized into two primary modules:

1. **Hospital Module:** This module provides demographic and baseline patient information. It includes essential details like age, gender, ethnicity, comorbidities, and diagnostic codes (ICD codes). This module serves as a foundation for cohort creation, allowing researchers to identify groups of patients based on various clinical characteristics, such as pre-existing conditions or disease severity at the time of admission.
2. **ICU Module:** The ICU module contains detailed, real-time clinical data during a patient's stay in the ICU. It includes data such as vital signs, lab results, treatments, medications, and specific events that occur in the ICU (e.g., mechanical ventilation, fluid resuscitation, or vasopressor administration). This module allows for an in-depth look at the physiological responses to treatment and interventions within the ICU setting.

Data Processing

In this section, we detail the comprehensive steps taken to prepare the final dataset for analysis, focusing on the identification of septic shock patients and their treatment with early vasopressors. The preprocessing workflow involves several key stages, including data cleaning, merging, imputation, and the creation of new variables. The objective is to structure the data in a way that allows us to evaluate the effect of early vasopressor administration on mortality outcomes in septic shock patients.

1. Data Sources and Initial Exploration

The dataset used in this analysis is from the MIMIC-IV database, which contains comprehensive data on patients admitted to intensive care units (ICU). This dataset includes variables related to patient demographics, diagnoses, treatment interventions, and outcomes. Key data files include:

- `icustays.csv.gz`: Contains information about ICU admissions, including timestamps for admission and discharge, and the type of care unit.
- `patients.csv.gz`: Contains patient demographic details such as age, gender, and race.
- `admissions.csv.gz`: Contains hospital admission details, including admission type, insurance information, and marital status.
- `diagnoses_icd.csv.gz`: Contains ICD codes for diagnoses, including codes for septic shock and related conditions.

We began by reading these files and performing basic exploratory data analysis (EDA) to understand the distribution of key variables, check for missing values, and verify data types. Initial checks revealed that the dataset had some missing values, especially in categorical variables such as insurance, marital status, and admission type.

Reasoning for Data Source Selection: The MIMIC-IV dataset is one of the most widely used resources for healthcare research, especially in the critical care domain. It provides a robust set of variables, including demographic, clinical, and outcome data, making it well-suited for analyzing sepsis and related treatment outcomes.

2. Creating the Sepsis Cohort

The second step involved defining the sepsis cohort, which focuses on patients diagnosed with septic shock according to Sepsis-3 criteria. We identified septic shock cases using ICD codes, including 78552, A41, R652, and others associated with sepsis and septic shock.

We merged this cohort with demographic data from the `patients.csv` file to form a comprehensive dataset that includes patient identifiers, age, gender, and other relevant sociodemographic factors. We also linked the sepsis cohort to ICU stay data to track patient progress in the ICU.

Reasoning for Cohort Selection: The focus on septic shock patients ensures that our analysis is directly relevant to the study's objectives. By using Sepsis-3 diagnostic criteria, we align with clinical standards in defining sepsis, ensuring that the findings are applicable in real-world medical settings.

3. Handling Missing Data

Handling missing data is a crucial step in healthcare data processing. In the MIMIC-IV dataset, missing values are common due to incomplete records, particularly in categorical variables such as insurance type and admission type, as well as continuous variables like Lactate and MAP.

- **Imputation for Categorical Variables:** For categorical variables like insurance and marital status, we used mode imputation (replacing missing values with the most frequent category). This method is appropriate when missingness is random, and the mode provides a reasonable estimate without significantly skewing the distribution.
- **Imputation for Continuous Variables:** For continuous variables like Lactate, MAP, and other clinical measurements, we applied KNN imputation (using 5 neighbors). KNN imputation was chosen because it leverages the relationships between features, allowing for more accurate imputation by considering similarities in patients' records. This method ensures that missing values are imputed based on similar patients with available data, preserving the relationships between clinical variables.

Reasoning for Imputation: Imputation allows us to retain as much of the dataset as possible, maintaining sample size and avoiding bias from dropping rows with missing data. In critical care datasets like MIMIC-IV, where each row represents an important patient's data, dropping rows with missing values would significantly reduce the available data, limiting the robustness of our analysis.

4. Filtering ICU Admissions

Filtering Based on Length of Stay (LOS): To ensure that we only include patients who had sufficient exposure to treatment and monitoring, we filtered out those with ICU stays shorter than 24 hours. These patients are unlikely to have received timely interventions, such as vasopressors, or to have experienced a full clinical course, making their data less relevant to the study.

Reasoning: Patients with very short ICU stays may not have been properly treated or diagnosed with septic shock, and their records may not be representative of the clinical course of septic shock. Excluding these cases ensures that our analysis focuses on patients who experienced the full trajectory of septic shock and its treatment.

5. Key Variables: Mean Arterial Pressure (MAP)

MAP Calculation:
Mean Arterial Pressure (MAP) was calculated using the standard formula:

$$\text{MAP} = \frac{(Systolic + 2 \times Diastolic)}{3}$$

Systolic and diastolic blood pressure measurements were extracted from the `chartevents.csv` dataset, and MAP was calculated to provide an essential measure of hemodynamic status in septic shock patients. **Reasoning:** MAP is a critical clinical variable used to assess the severity of shock and guide treatment decisions, particularly vasopressor administration. Including MAP as a derived variable allows for a more comprehensive analysis of shock severity and treatment response.

6. Identifying the Onset of Septic Shock

Onset Time for Septic Shock: We identified the onset of septic shock based on the first instance when MAP dropped below 65 mmHg, a widely accepted threshold for defining septic shock in clinical practice and when lactate values dropped below 2mmoL. The `charttime` of this first low MAP reading was recorded as the onset time for septic shock.

Reasoning: Accurately identifying the onset of septic shock is crucial for evaluating the impact of early vasopressor administration. By using MAP and lactate level as a marker, we align our approach with clinical standards, ensuring the findings are clinically relevant.

7. Merging Vasopressor Data

Vasopressor Data Integration:
We extracted norepinephrine administration data from the `inputevents.csv` file, where norepinephrine is recorded with itemid 221906. After pivoting the data to sum the total norepinephrine dose per patient, we merged this treatment data with the sepsis cohort to identify patients who received norepinephrine early (within 1 hour of septic shock onset).

Reasoning: Vasopressors are a key intervention in septic shock management. Early administration of norepinephrine is crucial for improving survival outcomes. Merging vasopressor data with the sepsis cohort allows us to define treatment groups (early vs. late administration), which is essential for assessing treatment effectiveness.

8. Feature Engineering: Creating Binary Flags

Early Vasopressor Administration Flag: We created a binary flag, `is_nephrohrine_given_within_1_hours`, to indicate whether norepinephrine was administered within the first hour of septic shock onset. This allows us to compare the outcomes of patients who received timely treatment versus those who did not.

Mortality Outcome Flag: We also created a flag, `dod_within_28_days_of_septic_shock`, to indicate whether a patient died within 28 days of septic shock onset. This was used as the primary outcome for assessing the effectiveness of early vasopressor treatment.

Reasoning:
Binary flags simplify analysis, especially in survival analysis or causal inference models. These flags help to clearly separate groups for comparison, enabling straightforward analysis of the impact of early vasopressor administration on mortality.

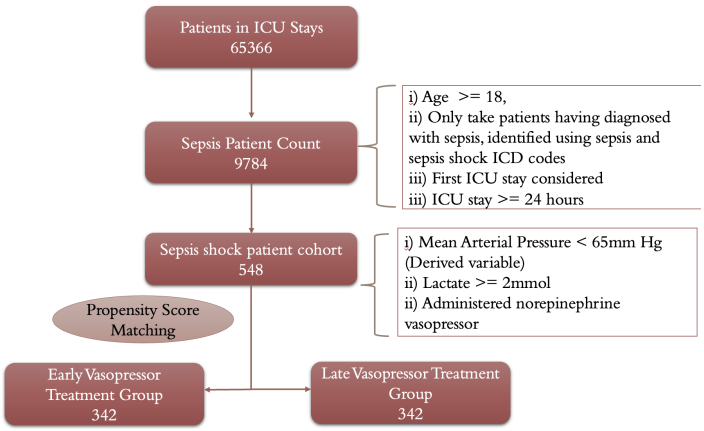


Figure 1: Patient Inclusion Criteria with count

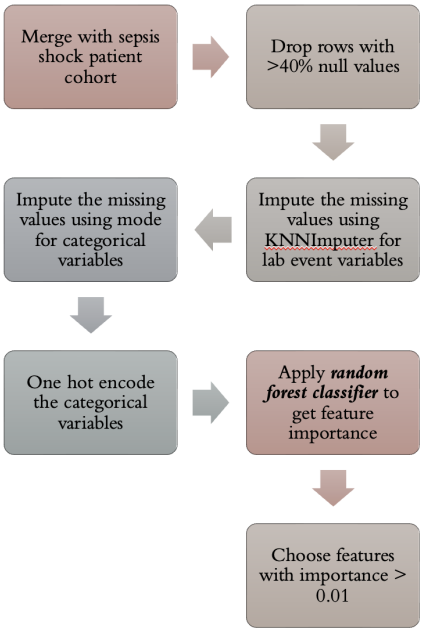


Figure 2: Brief Overview of Covariate Selection

Covariate/Feature Selection

In this section, we describe the process of selecting and analyzing covariates used in the study to evaluate the effects of early vasopressor administration in septic shock patients. We aimed to account for potential confounders that might influence both the treatment (vasopressor administration) and the outcome (28-day mortality).

1. Covariates Selection

The selected covariates fall into several broad categories:

- **Clinical variables:** These include key physiological measurements such as MAP, Lactate levels, and vital signs (e.g., Arterial Oxygen Saturation, pCO2, Temperature), which are critical indicators of shock severity and treatment needs.
- **Demographic variables:** These include Age, Gender, Race, and Marital Status, which are essential for controlling for sociodemographic biases.
- **Admission-related variables:** These include Insurance, Admission Type, and Length of Stay (LOS) in the ICU, all of which can influence treatment decisions and outcomes.

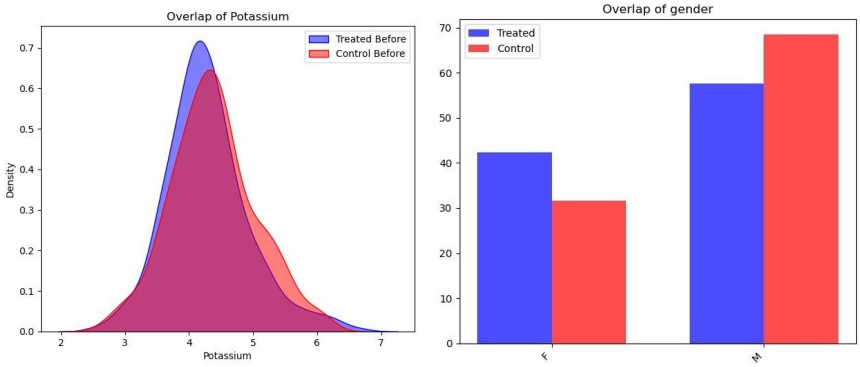
Reasoning for Covariate Selection: Covariates were selected based on their clinical relevance (e.g., MAP and Lactate are central to septic shock management) and their potential to influence both vasopressor administration and patient outcomes. Including a wide range of variables ensures that we control for as many confounders as possible and avoid bias in estimating the treatment effect. This comprehensive approach is in line with best practices in causal inference and healthcare research.

2. Analysis of Covariates Before Propensity Score Matching (PSM)

Before applying propensity score matching, we analyzed the distribution of covariates between the treated group (those who received early vasopressors) and the control group (those who did not). This analysis helps identify potential imbalances between the groups that could affect the validity of the treatment effect estimation.

- **Continuous Covariates (e.g., MAP, Lactate, Age):** We visualized the distribution of numerical covariates using Kernel Density Estimation (KDE) plots, which show the smooth distribution of the data. These plots provide insight into how well the treated and control groups overlap in terms of clinical characteristics.

Reasoning: KDE plots are effective in comparing distributions for continuous variables. By visualizing these distributions, we can assess if there is significant overlap between the treated and control groups. A lack of overlap might indicate the need for additional covariates or transformation to achieve balance after matching.



- **Categorical Covariates (e.g., Gender, Insurance, Admission Type):** For categorical variables, we used stacked bar plots to compare the proportions of each category between the treated and control groups. This allows us to observe any disparities in the distribution of categorical features, such as gender or race, that could affect the outcome.

Reasoning: Stacked bar plots provide a clear visual comparison of the proportions of each category in the treated vs. control groups. This helps to identify any potential imbalances in categorical covariates that could lead to confounding if not addressed.

3. Random Forest for Feature Selection

To identify the most important covariates influencing the likelihood of receiving early vasopressor treatment and the 28-day mortality outcome, we applied a Random Forest classifier to the pre-PSM data. Random Forest is a powerful ensemble learning method that helps identify important features by evaluating their contributions to prediction accuracy. The following covariates were included in the Random Forest model to predict both treatment (early vasopressor administration) and outcome (28-day mortality):

- **Clinical variables:** MAP, Lactate, Arterial Oxygen Saturation (SaO2), pCO2 (Arterial), Potassium, Temperature (Body Temperature)
- **Demographics:** Age, Gender, Race, Insurance, Marital Status
- **Admission-related variables:** Admission Type, Length of Stay (LOS), Last Care Unit

Techniques involved:

- **Grid Search:** We employed Grid Search to fine-tune the hyperparameters of the Random Forest model. This method systematically searches through multiple combinations of parameters (such as the number of trees, depth of trees, etc.) to identify the optimal configuration for feature selection.
- **Cross-Validation (CV):** We used 5-fold cross-validation to validate the model's performance, ensuring that the Random Forest model generalizes well to unseen data and isn't overfitting. This also helps provide reliable estimates of feature importance.

Reasoning for Random Forest: Random Forest was chosen because it handles complex, high-dimensional data and is effective at identifying non-linear relationships between variables. It can also handle both continuous and categorical data, making it ideal for this dataset, which includes a mix of clinical measures and demographic information.

- **Data Encoding:** For categorical variables, we used one-hot encoding to convert them into binary features. This allows the Random Forest algorithm to handle categorical data effectively, as it requires numerical input.

Reasoning for One-Hot Encoding: One-hot encoding is necessary to convert categorical variables into a format that can be processed by machine learning algorithms. It ensures that the model does not mistakenly interpret categorical variables as ordinal or continuous, which could distort the analysis.

- **Feature Importance:** After fitting the Random Forest model, we extracted the feature importance scores, which indicate the relative importance of each covariate in predicting treatment (early vasopressor administration) and outcome (mortality). This is an important step in understanding which factors are most predictive of treatment assignment and patient survival.

From the Random Forest analysis, we identified the top covariates with importance scores greater than or equal to a threshold of 0.012. The selected covariates were deemed most influential in predicting early vasopressor treatment and patient outcomes.

Feature	Importance
Temperature (Body Temperature)	0.098988
patientweight	0.091275
pO2 (Arterial)	0.088261
Potassium	0.087616
pCO2 (Arterial)	0.082281
pH (Arterial)	0.079033
MAP	0.078200
Arterial Oxygen Saturation (SaO2)	0.071906
Lactate	0.067854
Age	0.059636
marital_status_WIDOWED	0.018805
gender_M	0.014703

Figure 3: Features selected via Random Forest

Causal Inference and Propensity Score Matching (PSM) in Our Study

Causal inference is about understanding whether a treatment or intervention causes a specific outcome. Unlike correlation, which only identifies relationships between variables, causal inference aims to establish cause-and-effect relationships. In healthcare, this is vital for evaluating the impact of treatments like early vasopressor administration on outcomes such as mortality in septic shock. In ideal settings, we use Randomized Controlled Trials (RCTs) to draw causal conclusions because randomization ensures equal treatment across all groups. However, in real-world, observational data (like ours), there's often a risk of bias—patients who receive treatment may differ from those who do not. This is where Propensity Score Matching (PSM) comes in. PSM helps us estimate the treatment effect by matching treated and untreated individuals who are similar in all other respects (based on their covariates, like age or severity of illness). This technique aims to reduce selection bias, making the treated and control groups comparable, almost like in an RCT. In our study, we're using PSM to balance patients who received early vasopressors (the treated group) and those who did not (the control group). By matching based on factors like MAP, Lactate, and age, PSM allows us to compare these groups more fairly, ensuring that any differences in mortality are more likely due to the treatment itself, rather than other confounding factors. This approach helps us draw more accurate conclusions about the causal relationship between early vasopressor treatment and patient survival in septic shock.

Our Methodology and Corresponding Results

In this section, we describe the application of Propensity Score Matching (PSM) and the four different methods of causal inference used to estimate the effect of early vasopressor administration on 28-day mortality in septic shock patients. Each method offers a different way of estimating the treatment effect, and we provide detailed explanations of how they were implemented and their results.

1. Calculation of Propensity Scores

The propensity score is the predicted probability of receiving treatment (early vasopressor administration) based on the observed covariates. In our study, we used logistic regression to estimate the propensity scores, which are essentially the likelihood of a patient receiving vasopressors given their clinical characteristics.

Steps:

- **Logistic Regression Model:** A logistic regression model was used to predict the treatment indicator (`is_nephrohrine_given_within_1_hours`) based on the selected covariates. These covariates included clinical variables (e.g., MAP, Lactate, Age) and demographic variables (e.g., Gender, Race).
- **Logit Transformation:** We computed the logit of the propensity score for each individual. The logit is the logarithm of the odds of receiving treatment, which is the difference between the log of the probability of treatment and the log of the probability of not receiving treatment.

Evaluation scores for Logistic regression model:

```
Accuracy: 0.6788, ROC AUC: 0.7074, F1-score: 0.7684
```

pred	False	True
is_nephrohrine_given_within_1_hours		
0.0	80	126
1.0	50	292

Figure 4: Logistic regression evaluation scores

The propensity scores were visualized through Kernel Density Estimation (KDE) plots, which showed the overlap between the treated and control groups. This overlap is critical, as a lack of overlap would mean that the groups are not comparable, which would lead to biased estimates. **Results:** The KDE plots indicated that there was moderate overlap between the treated and control groups, which is essential for matching. However, the treatment group had a slightly higher concentration of patients with more severe conditions (as expected, due to the nature of early vasopressor treatment).

2. Propensity Score Matching

Propensity Score Matching Implementation

To balance baseline covariates between treated and untreated groups, we applied propensity score matching using a nearest-neighbor approach based on estimated logits (propensity scores). The matching process involved the following steps:

1. **Sorting:** Patients were sorted by propensity scores to facilitate efficient matching.
2. **Matching Algorithm:**
   - Each treated patient was matched to the closest untreated patient based on the difference in propensity scores.
   - Matches were identified both above and below the treated patient in the sorted list.
   - The closest untreated patient (minimizing the absolute distance in propensity scores) was selected.
3. **Validation:** We ensured a one-to-one match by assigning each untreated patient to at most one treated patient.
4. **Balancing Covariates:** After matching, balance in covariates was assessed to confirm comparability between groups.

This custom matching algorithm allowed for transparent, efficient, and reproducible identification of control units, facilitating robust causal inference.

Covariate Balance Analysis

To ensure the validity of causal inference, covariate balance between the treatment and control groups was evaluated both before and after matching. Descriptive statistics, including mean, standard deviation, quartiles, and frequencies, were computed for key variables, such as clinical measures (e.g., Potassium, pCO2) and categorical variables (e.g., marital status and gender).

Before Matching: Substantial differences were observed in the means and distributions of key covariates between the treatment and control groups, highlighting baseline imbalances. For instance, the mean Potassium level was 4.37 in the control group versus 4.28 in the treatment group, with similar imbalances for pCO2 (40.32 vs. 39.35). Likewise, categorical

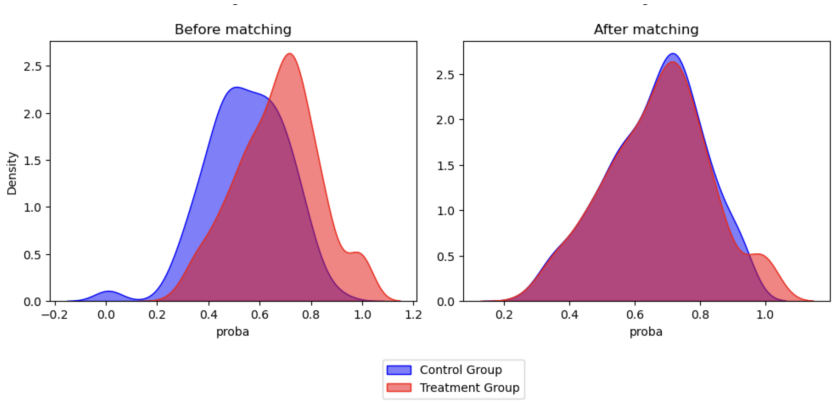


Figure 5: Propensity score and logit overlap before and after matching

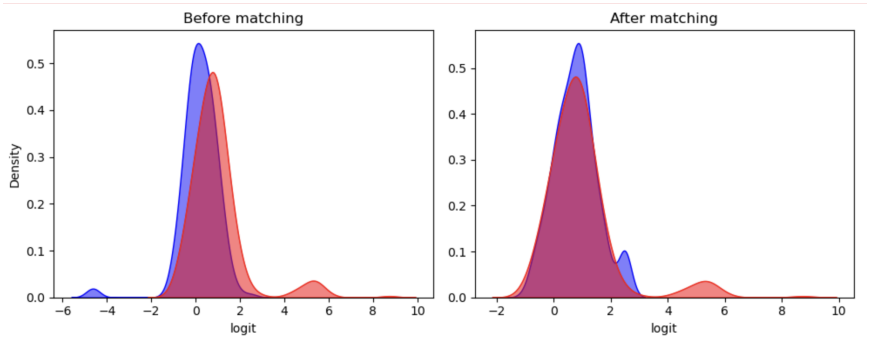


Figure 6: Logit overlap before and after matching

Continuous Variables

Potassium | Before matching

is_nephrohrine_given_within_1_hours	count	mean	std	min	25%	50%	75%	max
False	206.0	4.373041	0.667481	2.70	3.930833	4.314815	4.783333	6.10
True	342.0	4.277535	0.653258	2.55	3.865196	4.179808	4.600000	6.66

Potassium | After matching

is_nephrohrine_given_within_1_hours	count	mean	std	min	25%	50%	75%	max
False	342.0	4.307857	0.616905	2.70	3.939286	4.200000	4.66875	5.516667
True	342.0	4.277535	0.653258	2.55	3.865196	4.179808	4.60000	6.660000

pCO2 (Arterial) | Before matching

is_nephrohrine_given_within_1_hours	count	mean	std	min	25%	50%	75%	max
False	206.0	40.324425	7.186275	21.111111	35.552083	40.000000	43.720395	64.111111
True	342.0	39.346165	8.236149	16.375000	34.355615	38.666667	43.637080	79.027027

pCO2 (Arterial) | After matching

is_nephrohrine_given_within_1_hours	count	mean	std	min	25%	50%	75%	max
False	342.0	39.304737	7.152834	21.111111	34.888889	39.672414	43.789474	64.111111
True	342.0	39.346165	8.236149	16.375000	34.355615	38.666667	43.637080	79.027027

Categorical Variables

marital_status_WIDOWED | Before matching

is_nephrohrine_given_within_1_hours	count	unique	top	freq
False	206	2	False	182
True	342	2	False	321

marital_status_WIDOWED | After matching

is_nephrohrine_given_within_1_hours	count	unique	top	freq
False	342	2	False	316
True	342	2	False	321

gender_M | Before matching

is_nephrohrine_given_within_1_hours	count	unique	top	freq
False	206	2	True	141
True	342	2	True	197

gender_M | After matching

is_nephrohrine_given_within_1_hours	count	unique	top	freq
False	342	2	True	200
True	342	2	True	197

Figure 7: Covariates In Treatment And Control Group Are Closer Than Before

variables such as gender and marital_status showed uneven frequencies, with 68.4% males in the control group compared to 57.6% in the treatment group.

After Matching: The matching process significantly reduced these differences. For example, the mean Potassium levels in the treatment and matched control groups were closely aligned at 4.28 and 4.31, respectively, with comparable standard deviations of 0.65 and 0.62. Similarly, the mean pCO2 values were balanced between groups, ensuring distributions were more comparable. The balance in categorical variables also improved. For marital_status_WIDOWED, the proportion of widowed patients in the control group was closely matched to the treated group post-matching (26 vs. 21). The frequencies for gender_M were also well-aligned, with 58.5% males in the control group compared to 57.6% in the treated group.

These improvements demonstrate that the matching process effectively minimized confounding by aligning the treatment and control groups on key covariates. This ensures that any observed differences in outcomes can be more confidently attributed to the treatment effect rather than baseline imbalances, facilitating robust causal inference.

3. Causal Inference Methods Using Propensity Score Matching

After calculating the propensity scores and balancing the groups, we used four causal inference methods to estimate the treatment effect. Each method provides a different perspective on the treatment effect, but all rely on the propensity scores and the matched data.

Methods

Method 1: Average Treatment Effect on the Treated (ATT)

The ATT measures the treatment effect for individuals who received the treatment, comparing outcomes between treated and matched control groups. After matching treated units to controls based on propensity scores, the difference in mean mortality outcomes between the two groups is computed.

Method 2: Inverse Probability Weighting (IPW)

IPW estimates the Average Treatment Effect (ATE) by adjusting for confounding through reweighting observations based on their propensity scores. The dataset is split into treated and control groups. For each observation, a weight is computed inversely proportional to the propensity score: for treated units, the weight is $1/P(T = 1|X)$, and for controls, the weight is $1/(1 - P(T = 1|X))$. The weighted averages of outcomes for treated and control groups are then calculated, and the ATE is derived as the difference between these averages. This method mimics a randomized experiment, adjusting for observed covariates.

Method 3: Causal Effect Estimation with Matching and Covariate Adjustment

This method estimates the causal effect of treatment on an outcome variable by matching treated and control units with similar covariate profiles, followed by covariate adjustment. The outcome and treatment indicators are extracted from the matched dataset, and covariates are standardized to ensure comparability. A causal model is created and the treatment effect is estimated by matching treated and control pairs, minimizing baseline differences. This approach provides robust estimates of the treatment effect by simulating randomized trial conditions.

Method 4: Causal Effect Estimation Using DoWhy Framework

The DoWhy framework estimates the causal effect of early vasopressor treatment on 28-day mortality in septic shock. A causal model is initialized with treatment, outcome, and confounding variables. The causal graph is visualized, and the causal effect is identified using the backdoor criterion to account for confounding. Propensity score weighting is used to estimate the treatment effect, and the estimate is validated using robustness checks, including placebo treatments. The method provides transparency in causal assumptions and enhances confidence in the results through validation techniques.

Results and Conclusion

Results

Four methods were used to estimate the causal effect of administering vasopressors within 1 hour on 28-day mortality in septic shock patients. The results are summarized as follows:

- **Method 1: Average Treatment Effect on the Treated (ATT)**: The ATT was estimated as -0.0439, indicating a 4.39% reduction in 28-day mortality for treated patients compared to matched controls.
- **Method 2: Inverse Probability Weighting (IPW)**:
  - Weighted outcomes:
    - Treated group: 0.6546
    - Control group: 0.7421
  - The ATE was -0.0875, suggesting an 8.75% reduction in mortality.
- **Method 3: CausalModel Package**: The estimated ATE was -0.064, reflecting a 6.4% reduction in 28-day mortality.
- **Method 4: DoWhy Framework**: The DoWhy analysis estimated a 2.65% reduction in 28-day mortality for patients treated within 1 hour.

Conclusion

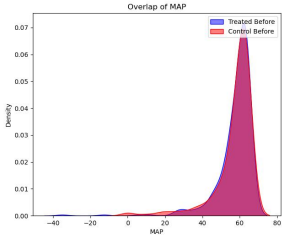
Across all methods, the results consistently indicate that early administration of vasopressors is associated with a reduction in 28-day mortality among septic shock patients. While the magnitude of the effect varies slightly across methods due to differences in statistical techniques, the overall conclusion is robust: timely intervention with vasopressors likely improves survival outcomes. The consistency of findings across multiple methodologies enhances the reliability of the conclusion and underscores the importance of prompt treatment decisions in clinical settings.

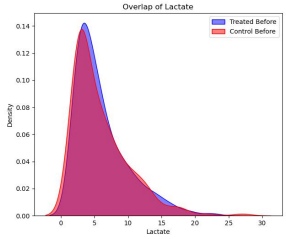
Future Work

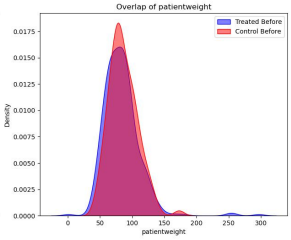
Future research can focus on assessing the long-term effects of vasopressor administration by analyzing outcomes beyond 28-day mortality, such as 90-day survival rates, quality of life post-discharge, and long-term organ function recovery. Incorporating survival analysis techniques, including Kaplan-Meier survival curves and Cox proportional hazards models, can provide insights into the sustained benefits or potential risks associated with early vasopressor use. Additionally, the analysis can be expanded to evaluate the efficacy of different types of vasopressors, such as dopamine and vasopressin, and their timing in septic shock treatment. This would help identify optimal treatment protocols tailored to patient subgroups. Future studies could also investigate the effects of other ICU interventions, such as fluid resuscitation strategies or corticosteroid administration, and their interaction with vasopressor timing to optimize overall treatment strategies for septic shock.

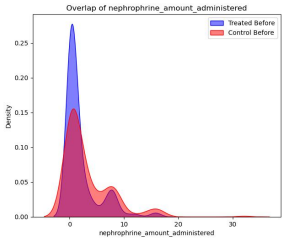
APPENDIX

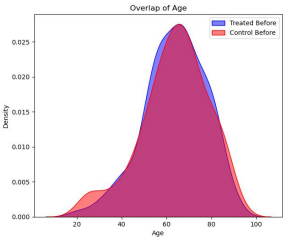
Analysis of covariates before random forest feature selection

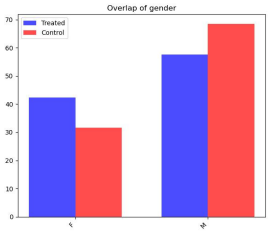


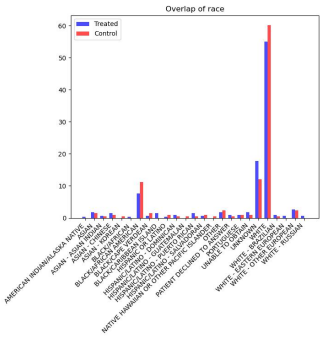


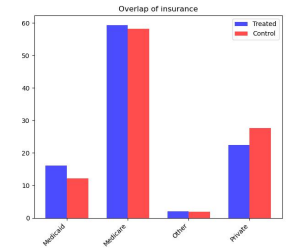


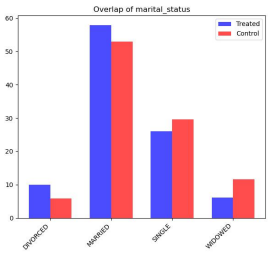


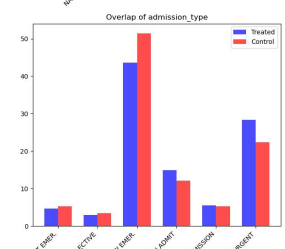


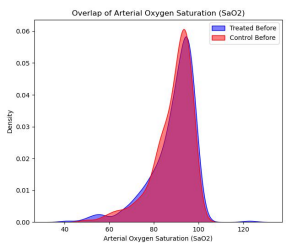


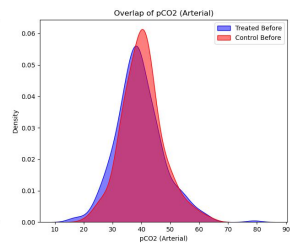


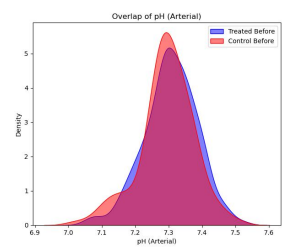


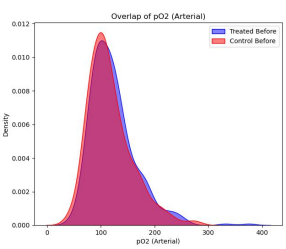


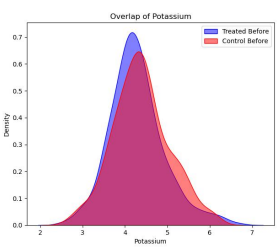


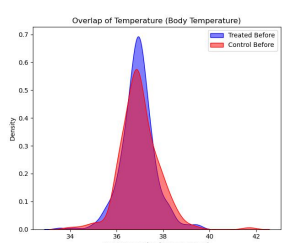




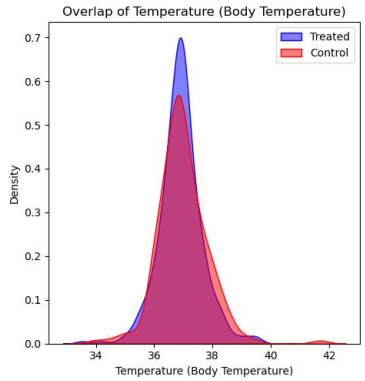


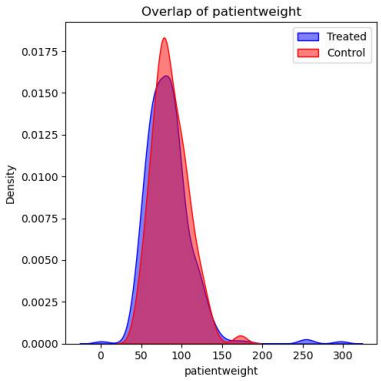


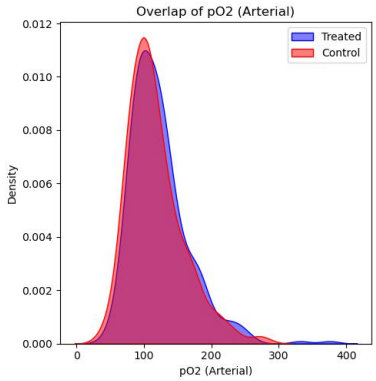


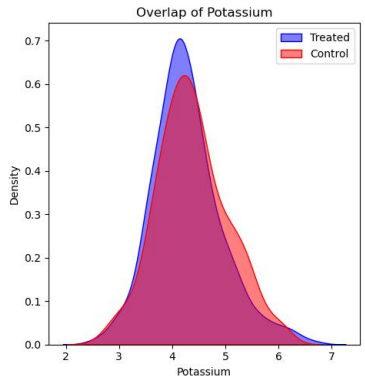


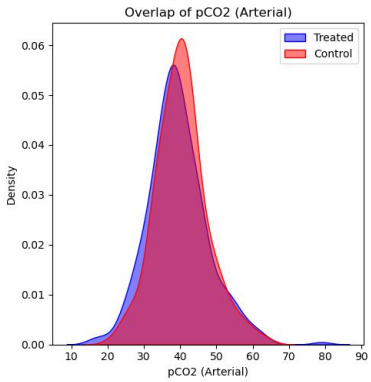
Analysis of covariates after random forest feature selection

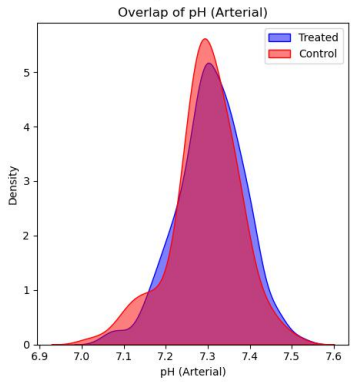


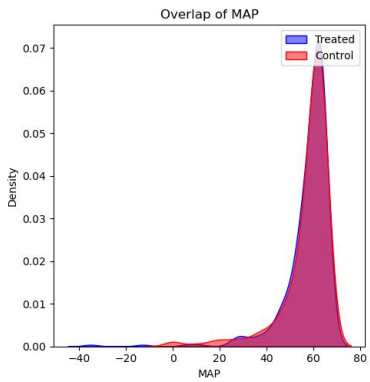


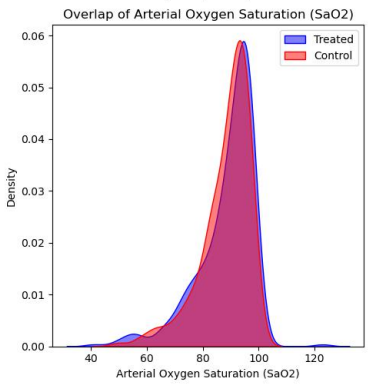


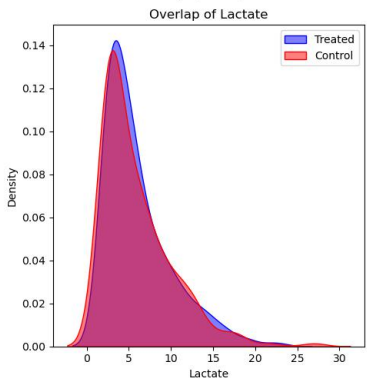


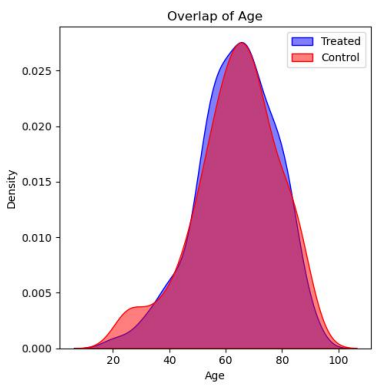


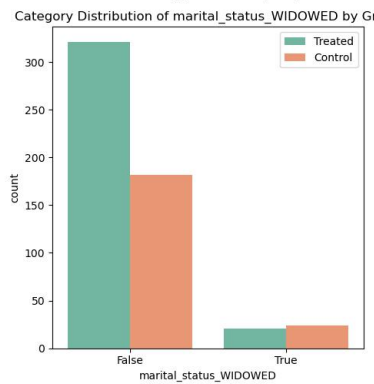


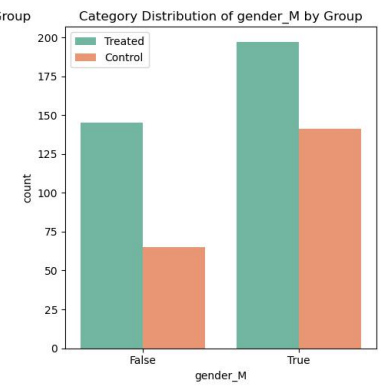


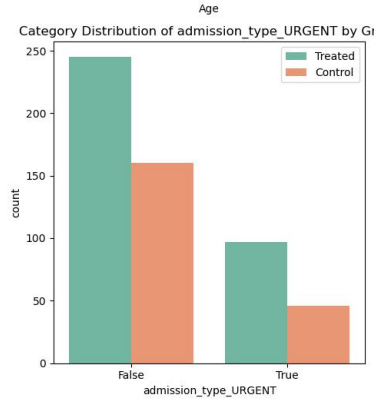


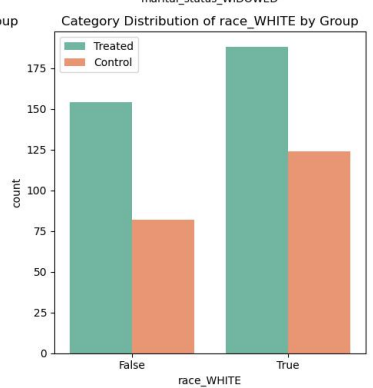












Analysis of covariates after and before matching