OrthoMortPred: In-Hospital Mortality Prediction for Orthopedic Patients

Carvalho Fr1,3, Camacho AB1,2

*1*Faculty of Medicine and Biomedical Sciences, University of Algarve*, Faro, Portugal*

2 Central Lisbon University Hospital Center, CRI-Orthopedic Traumatology

*3Centre of Marine Sciences (CCMAR/CIMAR LA), University of Algarve, Faro, Portugal*

*University of Algarve - Campus de Gambelas, 8005-139 Faro, Portugal*

*Corresponding author: Filipe Ricardo Carvalho* ***[frcarvalho@ualg.pt](mailto:frcarvalho@ualg.pt)***

Abstract

In-hospital mortality among orthopedic patients remains a significant concern in healthcare systems worldwide. This study aimed to develop and validate a clinical prediction model for in-hospital mortality risk in orthopedic patients using machine learning techniques. We utilized a dataset from the Central Lisbon University Hospital Center's CRI-Orthopedic Traumatology department, comprising 3,132 patients from 2021 to 2023. After extensive data preprocessing and feature selection, we employed various machine learning algorithms, with LightGBM Classifier emerging as the best-performing model. The final model, using 11 selected features, achieved an overall accuracy of 93% and an area under the ROC curve of 0.93. The most influential predictors were emergency admission date and time, age, and pre-operative days. Cross-validation results supported the model's robustness, with scores ranging from 0.97 to 0.99 across five folds. SHAP (SHapley Additive exPlanations) analysis provided insights into the model's decision-making process, confirming the importance of temporal factors and patient demographics in predicting mortality risk. The model demonstrated high precision and recall for both survival (0.94 and 0.98) and mortality (0.91 and 0.78) predictions. This study presents a highly accurate and interpretable model for predicting in-hospital mortality in orthopedic patients. By leveraging readily available clinical data, this tool has the potential to enhance risk stratification, inform clinical decision-making, and ultimately improve patient outcomes. Future research should focus on external validation and assessing the model's impact on clinical practice when implemented in diverse healthcare settings.

Introduction

Predicting the risk of in-hospital decease is a critical aspect of patient care across various medical disciplines. Early identification of high-risk patients can inform clinical decision-making, resource allocation, and targeted interventions to improve patient outcomes (1). In recent years, disease prediction models have increasingly incorporated machine learning techniques, nomograms, and electronic health record data. These approaches are now recognized as powerful tools for assessing the predictive value of clinical models (2,3). While numerous studies have explored predictive models for in-hospital mortality in various medical settings (4–6), such as infectious diseases (4,7), COVID-19 (8), cardiovascular diseases(3,6,9,10), and intensive care units(9,11), there is a paucity of research specifically focused on the orthopaedic department(12,13). Orthopaedic patients often present with unique challenges and comorbidities, including musculoskeletal injuries, degenerative conditions, and postoperative complications(12). These factors can contribute to increased risk of in-hospital mortality, necessitating tailored predictive models that account for the specific characteristics and risk factors prevalent in this patient population. The development of a robust and accurate prediction model for in-hospital decease in the orthopaedic department could have significant implications for patient care (12). Early identification of high-risk patients could prompt timely interventions, such as enhanced monitoring, targeted treatment strategies, or referral to higher levels of care. Additionally, such a model could aid in resource allocation, ensuring that appropriate medical resources are directed toward high-risk patients, potentially improving overall patient outcomes and optimizing healthcare resource utilization (14). This research aims to address the gap in predicting in-hospital mortality specific to the orthopaedic department by developing and validating a predictive model tailored to this patient population. By leveraging relevant clinical data and advanced modelling techniques, this study seeks to provide a valuable tool to aid clinicians in identifying high-risk orthopaedic patients and implementing appropriate measures prior to surgery to mitigate the risk of in-hospital decease.

Methods

1. Study Design and Data Source

This retrospective cohort study utilized data from the "CRI\_Orthopedic\_Traumatology\_21\_22\_23.csv" dataset, which contains records from the Central Lisbon University Hospital Center's CRI-Orthopedic Traumatology department spanning from 2021 to 2023. The dataset encompasses a comprehensive range of variables related to orthopaedic surgical procedures, including patient demographics, surgical scheduling, intervention details, medical specialties, operating theatre information, patient flow, diagnostic and procedural data, anaesthesia information, surgical team composition, operational timings, and billing status.

1. Cohort Characteristics

The study population comprised 3,132 individuals, consisting of 1,960 males (62.6%) with a mean age of 76.33 years (SD = 17.64 years) and 1,172 females (37.4%) with a mean age of 57.35 years (SD = 21.53 years). In-hospital mortality was observed in 161 males (8.2% of the male) and 85 females (7.3% of the female) (figure 1).

1. Outcome Variable Definition

The primary outcome of interest, defined as the target variable, was the risk of in-hospital decease. This variable was operationalized to include:

1. All cases of in-hospital mortality.
2. All patients with age belonging to the 83th percentile or above.

The inclusion of the latter group in our target variable is justified by the observation that this age group accounted for 19% of all decease cases, representing the highest risk factor among all variables analysed (Figure 2).

1. Data Preprocessing and Quality Assurance

Extensive data cleaning and preprocessing steps were performed to handle missing values and prepare the data for analysis. We utilized pandas for data manipulation and preprocessing. Columns with almost 100% null values were identified and removed. Columns with more than 50% null values were manually curated and nan were filled with zeros, while columns with 50% to 20% null values were filled with either zeros for object columns or the mean value for numerical columns. For columns with 20% to 1% null values, we used the KNNImputer from scikit-learn to impute missing values in numerical columns, and the SimpleImputer was used to impute missing values in categorical columns using the most frequent value (Figure 4).

1. Feature Correlation Analysis

To understand which variables are more related to decease, we plotted the correlation of the different variables with the decease variable using seaborn and matplotlib for visualization. We considered variables of interest those that presented a correlation higher or lower than 0.1 or -0.1 (Figure 4).

1. Machine Learning Model Development and Evaluation

To ensure a robust evaluation, the data was split into training and test sets with a 70:30 split ratio, utilizing sklearn's train\_test\_split function with a fixed random state for reproducibility. Furthermore, the SMOTE (Synthetic Minority Over-sampling Technique) algorithm from imblearn package was applied to balance the class distribution in the training set, and data was normalized using sklearn's MinMaxScaler function (Figure 4).

6.1. Feature Selection and Algorithm Comparison

To determine the optimal estimator for feature selection and the best classifier, we implemented a comprehensive approach using sklearn. For feature selection, we utilized multiple techniques, including Recursive Feature Elimination with Random Forest and SelectFromModel with various base estimators such as Random Forest, Gradient Boosting, Logistic Regression, Extra Trees, and AdaBoost. We then evaluated a diverse set of classification algorithms, including Random Forest, Gradient Boosting, Support Vector Machine, Decision Tree, XGBoost (from xgboost library), LightGBM (from lightgbm library), CatBoost (from catboost library), Extra Trees, AdaBoost, and Ridge Classifier. For each combination of feature selector and classifier, we constructed a pipeline using sklearn that first applies the feature selection method and then trains the classifier on the selected features (Figure 4).

6.2. Hyperparameter Optimization

To optimize the performance of the LightGBM Classifier, a grid search approach was adopted using sklearn's GridSearchCV. The hyperparameter grid included various parameters for both feature selection and classification. The grid search was performed using 5-fold cross-validation, and the ROC-AUC score was employed as the evaluation metric (Figure 4).

6.3. Ensemble Modeling: Stacking Approach

After identifying the best estimator and classifier, we proceeded to develop a stacking model using sklearn's StackingClassifier. This ensemble approach combined base learners, such as Gradient Boosting Classifier, Decision Tree Classifier, XGBoost Classifier, and CatBoost Classifier, with a LightGBM Classifier serving as the meta-classifier (Figure 4).

6.4. Performance Metrics and Cross-Validation

The study employed a comprehensive set of evaluation metrics from sklearn to assess the performance of the developed models. These metrics included the classification report, confusion matrix, and ROC-AUC score. Additionally, 5-fold cross-validation scores were computed to ensure the robustness and generalizability of the models. We used matplotlib and seaborn for visualizing these evaluation metrics.

1. Model Interpretation and Feature Importance Analysis

To gain insights into the model's decision-making process, we employed SHAP (SHapley Additive exPlanations) values. The shap library was used to compute and visualize the impact of each feature on the model's predictions, providing a deeper understanding of the factors influencing the risk of in-hospital decease.

1. **Ethical Considerations**

The study protocol was approved by the Ethics Committee of the Central Lisbon University Hospital Center. Given the retrospective nature of the study using de-identified data, the requirement for individual patient consent was waived. All patient data were anonymized and handled in compliance with the General Data Protection Regulation (GDPR) to ensure patient privacy and confidentiality.The researchers had no access to identifiable patient information during the data analysis process. The study was designed and conducted with the primary aim of improving patient care and outcomes in orthopedic settings, adhering to the principles of beneficence and non-maleficence.

1. **Availability of Resources**

To ensure transparency and reproducibility, all code used for data preprocessing, analysis, and model development, as well as the anonymized dataset, are available on GitHub at the following repository: [insert GitHub repository URL].

This repository includes:

* The complete Python scripts used for data cleaning, feature engineering, and model development
* Documentation on the data structure and variable definitions
* Instructions for replicating the analysis
* The final trained model (excluding any patient-identifiable information)

We encourage other researchers to utilize these resources for validation studies or further development of predictive models in orthopedic care. Please note that while the data has been anonymized, users should still handle it responsibly and in accordance with ethical guidelines for medical research data.

For any questions regarding the use of these resources or requests for additional information, please contact the corresponding author.

Results

The study aimed to develop and validate a clinical prediction model for in-hospital mortality among orthopaedic patients, addressing a significant gap in the literature. Using a dataset from the Central Lisbon University Hospital Center's CRI-Orthopedic Traumatology department, the research team analysed records of 3,132 patients from 2021 to 2023.

The cohort characteristics revealed a gender distribution of 62.6% males and 37.4% females, with mean ages of 76.33 and 57.35 years, respectively. In-hospital mortality rates were similar between genders, with 8.2% for males and 7.3% for females (Figure 3). In the feature correlation analysis, age emerged as the strongest predictor of in-hospital mortality, with a correlation of 0.22. Other variables showing correlations included anaesthesia risk (0.21), GDH level (0.18), and diagnostic code (0.18), suggesting that these factors also play important roles in predicting mortality.The machine learning model development process involved rigorous feature selection and algorithm comparison. After implementing a pipeline that involved different algorithms as selectors and different algorithms as classifiers, the combination of selector and classifier that produced the best results in this pipeline was the combination of GradientBoostingClassifier and LGBMClassifier. This combination was chosen for the subsequent analyses. The process of choosing the number of features to be selected and sent to the classifier was done manually, which involved many experiments to determine the best number of features to provide to the classifier. After several attempts, the number of max\_features was set to 11. After determining the optimal number of features, we conducted an extensive grid search. This process involved testing a vast array of hyperparameter combinations to fine-tune our model. The grid search explored various options for key parameters of our chosen algorithm. After a comprehensive evaluation, the best-performing configuration was identified with the following hyperparameters: learning\_rate: 0.8, max\_depth: 6, n\_estimators: 20, num\_leaves: 20. This combination of hyperparameters resulted in the most effective model performance based on our evaluation metrics. The stacking model, which combined multiple base learners including GradientBoostingClassifier, DecisionTreeClassifier, XGBClassifier, and CatBoostClassifier, with LGBMClassifier as the meta-classifier, was implemented to potentially improve predictive performance. The base learners were designed to capture different aspects of the data, while the meta-classifier was meant to learn from their collective outputs. Despite this sophisticated approach, the stacking model did not outperform the simpler combination of the feature selector (SelectFromModel with GradientBoostingClassifier) and the standalone LGBMClassifier. The LightGBM Classifier, used as an individual model, emerged as the best-performing model, demonstrating high accuracy and stability. The model achieved an overall accuracy of 0.93, with precision and recall of 0.94 and 0.98 for survival (class 0), and 0.91 and 0.78 for mortality (class 1), respectively. These metrics indicate a strong predictive performance, particularly in identifying patients likely to survive. The confusion matrix provided a detailed breakdown of the model's predictions, correctly identifying 700 survival cases and 175 mortality cases, while misclassifying 17 cases as false positives and 48 as false negatives.

Cross-validation results further supported the model's robustness, with scores ranging from 0.97 to 0.99 across five folds, with a mean CV score of 0.99. This high score suggests that the model's performance is consistent across different subsets of the data, enhancing its generalizability to new patients.

Feature importance represents the relative contribution of each feature to the performance of a machine learning model. The importance values indicate how influential each feature is for the model's predictions. 'EMERGENCY ADMISSION DATE TIME' with a score of 87 was the most important feature, in the prediction of high and low risk of decease meaning that splits using this feature resulted in substantial improvements. 'AGE' had a score of 82, indicating it is also highly influential in our model prediction. 'PRE OP DAYS' add a score of 54, showing it´s crucial for the model. The importance score reflects the cumulative contribution of these splits across all trees in the ensemble. SHAP (SHapley Additive exPlanations) values offer in-depth insights into model predictions by quantifying the impact of each feature. The summary plot visualizes this impact, with the color representing the feature value and the x-axis displaying the SHAP value (Figure 8.B). For instance, the feature “AGE” shows the highest mean SHAP value, highlighting its significant influence on the model’s predictions. Figure 8.A illustrates the SHAP values for a specific instance, revealing how individual features contribute to that particular prediction. This method aids in interpreting and validating model behavior. In the given example, a 105-year-old patient’s high SHAP value for “AGE” indicates that this feature has a strong effect on the prediction of a high risk of death.

Discussion

The present study aimed to develop and validate a clinical prediction model for in-hospital mortality among orthopedic patients, addressing a significant gap in the literature. Our findings demonstrate that machine learning techniques (15), specifically the LightGBM Classifier, can effectively predict the risk of in-hospital decease using a set of readily available clinical variables (10). The model achieved an overall accuracy of 93%, with an area under the ROC curve of 0.93, indicating excellent discriminative ability and can be used to inform the clinicians of the risk of decease in the moment prior to surgery. This information can anticipate and contribute to a better management of high-risk patients, allowing for more tailored care strategies and resource allocation (16).

This proactive approach may lead to improved patient outcomes, reduced complications, and potentially lower mortality rates (8). One of the most striking findings of our study is the paramount importance of emergency admission date and time as the strongest predictor of mortality risk. This result suggests that the timing of hospital admission may play a crucial role in patient outcomes, possibly reflecting variations in hospital resources, staffing levels, or the severity of the patient's condition upon arrival (17). The significance of this temporal factor warrants further investigation and may have important implications for hospital resource allocation and staffing strategies(18).

Age emerged as the second most important predictor, corroborating extensive previous research on the impact of advanced age on orthopaedic outcomes (19). This finding underscores the need for heightened vigilance and potentially more aggressive management strategies in elderly orthopaedic patients. The significance of age as the second most important predictor is primarily since most deaths in our study occur in patients with age above the 50th percentile. Notably, in the 83rd percentile or higher, the mortality rate reaches 20% among all patients, which motivated the inclusion of this group (percentile > 83) together with decease patients in our target population. Interestingly, the number of pre-operative days also proved to be a strong predictor of mortality risk. This suggests that delays in surgical intervention may adversely affect patient outcomes and increase mortality (20,21), highlighting the potential benefits of early surgical management in high-risk patients. The clinical implications of our findings are substantial. The developed model has the potential to assist physicians in early identification of high-risk orthopaedic patients, allowing for timely implementation of targeted interventions. These might include enhanced monitoring, pre-operative optimization, or consultation with additional specialists. Moreover, the identification of pre-operative days as a significant risk factor suggests that strategies to reduce surgical delays could potentially improve patient outcomes.

Despite its strengths, our study has limitations that warrant consideration. First, as a single-centre study, the generalizability of our findings to other institutions or healthcare systems may be limited. Second, while we employed advanced imputation techniques to handle missing data, this approach may have introduced some bias into our results. Third, our definition of in-hospital death risk, which included patients in the 83rd percentile of age or above, may have influenced the results and requires further validation.

The significance of the PROCEDURE DESCRIPTION, DRG code, BASE PROCEDURE ORDER DESCRIPTION in predicting mortality risk suggests that certain diagnostic categories may be associated with higher risk. For example the procedure OQH734Z as a mortality rate of 26%. Future studies could delve deeper into these associations to identify specific patient subgroups that may benefit from more intensive management or novel interventions.

The use of SHAP (SHapley Additive exPlanations) values in our analysis provides valuable insights into the model's decision-making process. This approach allows for a more transparent and interpretable machine learning model (22), which is crucial for building trust and facilitating adoption in clinical settings. The SHAP analysis confirmed the importance of age and emergency admission timing, while also highlighting the role of other factors such as anaesthesia risk and diagnostic codes in predicting mortality risk.

External validation of this model in diverse populations and institutions is essential to confirm its generalizability (23). Prospective studies evaluating the impact of implementing this prediction model in clinical practice would be valuable in assessing its real-world utility. Additionally, further investigation into the relationship between emergency admission timing and patient outcomes could provide insights into potential targets for quality improvement initiatives.

It is important to note that while our model demonstrates high predictive accuracy, it should be viewed as a tool to augment, rather than replace, clinical judgment (24). The complex interplay of factors contributing to in-hospital mortality in orthopaedic patients necessitates a nuanced approach to patient care that combines predictive analytics with experienced clinical decision-making.

1. Housseine N, Rijken MJ, Weller K, Nassor NH, Gbenga K, Dodd C, et al. Development of a clinical prediction model for perinatal deaths in low resource settings. EClinicalMedicine. 2022;44.

2. Xie W, Li Y, Meng X, Zhao M. Machine learning prediction models and nomogram to predict the risk of in-hospital death for severe DKA: A clinical study based on MIMIC-IV, eICU databases, and a college hospital ICU. Int J Med Inform. 2023;174.

3. Yan M, Liu H, Xu Q, Yu S, Tang K, Xie Y. Development and validation of a prediction model for in-hospital death in patients with heart failure and atrial fibrillation. BMC Cardiovasc Disord. 2023;23(1).

4. Lowlaavar N, Larson CP, Kumbakumba E, Zhou G, Ansermino JM, Singer J, et al. Pediatric in-hospital death from infectious disease in Uganda: Derivation of clinical prediction models. PLoS One. 2016;11(3).

5. Ewig S, Bauer T, Richter K, Szenscenyi J, Heller G, Strauss R, et al. Prediction of in-hospital death from community-acquired pneumonia by varying crb-age groups. European Respiratory Journal. 2013;41(4).

6. Richardson M, Howell N, Freemantle N, Bridgewater B, Pagano D. Prediction of in-hospital death following aortic valve replacement: A new accurate model. European Journal of Cardio-thoracic Surgery. 2013;43(4).

7. Gonçalves BP, Procter SR, Paul P, Chandna J, Lewin A, Seedat F, et al. Group B streptococcus infection during pregnancy and infancy: estimates of regional and global burden. Lancet Glob Health. 2022 Jun 1;10(6):e807–19.

8. Garrafa E, Vezzoli M, Ravanelli M, Farina D, Borghesi A, Calza S, et al. Early prediction of in-hospital death of covid-19 patients: A machine-learning model based on age, blood analyses, and chest x-ray score. Elife. 2021;10.

9. Garg NK, Ray S, Mathur A. Abstract 10932: Prediction of 30-Day Hospital Readmission in High-Risk Atherosclerotic Cardiovascular Disease Patients Using Machine Learning Methods on Electronic Health Record Data from Medical Information Mart for Intensive Care-3 Database. Circulation. 2021;144(Suppl\_1).

10. Zheng C, Tian J, Wang K, Han L, Yang H, Ren J, et al. Time-to-event prediction analysis of patients with chronic heart failure comorbid with atrial fibrillation: a LightGBM model. BMC Cardiovasc Disord. 2021;21(1).

11. Yu L, Zhou X, Li Y, Liu M. Establishment and evaluation of early in-hospital death prediction model for patients with acute pancreatitis in intensive care unit. Chinese Critical Care Medicine. 2023;35(8).

12. Etscheidt J, McHugh M, Wu J, Cowen ME, Goulet J, Hake M. Validation of a prospective mortality prediction score for hip fracture patients. European Journal of Orthopaedic Surgery and Traumatology. 2021;31(3).

13. Cabrera A, Bouterse A, Nelson M, Dietrich C, Razzouk J, Oyoyo U, et al. Prediction of In-Hospital Mortality Following Vertebral Fracture Fixation in Patients With Ankylosing Spondylitis or Diffuse Idiopathic Skeletal Hyperostosis: Machine Learning Analysis. Int J Spine Surg. 2024;18(1).

14. Chen J, Bai Y, Liu H, Qin M, Guo Z. Prediction of in-hospital death following acute type A aortic dissection. Front Public Health. 2023;11.

15. Huang G, Liu H, Gong S, Ge Y. Survival Prediction After Transarterial Chemoembolization for Hepatocellular Carcinoma: a Deep Multitask Survival Analysis Approach. J Healthc Inform Res. 2023;7(3).

16. Hamsen U, Drotleff N, Lefering R, Gerstmeyer J, Schildhauer TA, Waydhas C. Mortality in severely injured patients: Nearly one of five non-survivors have been already discharged alive from ICU. BMC Anesthesiol. 2020;20(1).

17. Guttmann A, Schull MJ, Vermeulen MJ, Stukel TA. Association between waiting times and short term mortality and hospital admission after departure from emergency department: Population based cohort study from Ontario, Canada. BMJ. 2011;342(7809).

18. Valli G, Galati E, De Marco F, Bucci C, Fratini P, Cennamo E, et al. In-hospital mortality in the emergency department: clinical and etiological differences between early and late deaths among patients awaiting admission. Clin Exp Emerg Med. 2021;8(4).

19. Walicka M, Puzianowska-Kuznicka M, Chlebus M, Śliwczyński A, Brzozowska M, Rutkowski D, et al. Relationship between age and in-hospital mortality during 15,345,025 non-surgical hospitalizations. Archives of Medical Science. 2021;17(1).

20. Jones S, Moulton C, Swift S, Molyneux P, Black S, Mason N, et al. Association between delays to patient admission from the emergency department and all-cause 30-day mortality. Emergency Medicine Journal. 2022;

21. Zafar MH, Zaka Ur Rehman TA, Khan MS, Ahmed S, Shariff A. The Impact of Delayed Surgical Care on Patient Outcomes With Alimentary Tract Perforation: Insight From a Low-Middle Income Country. Cureus. 2022;

22. Bosschieter TM, Xu Z, Lan H, Lengerich BJ, Nori H, Painter I, et al. Interpretable Predictive Models to Understand Risk Factors for Maternal and Fetal Outcomes. J Healthc Inform Res. 2024;8(1).

23. Yang CC. Explainable Artificial Intelligence for Predictive Modeling in Healthcare. J Healthc Inform Res. 2022;6(2).

24. Zeng Z, Yao L, Roy A, Li X, Espino S, Clare SE, et al. Identifying Breast Cancer Distant Recurrences from Electronic Health Records Using Machine Learning. J Healthc Inform Res. 2019;3(3).

A graph of a number of numbers and a chart of a number of numbers

Description automatically generated with medium confidence

**Figure 1. Distribution of Age by Risk of Decease Category.** Violin plot shows the age distribution across risk categories(0 low risk, 1 high risk), with inner boxplots indicating central measures. Descriptive statistics are annotated above each category.

A graph of cases and number of diseases

Description automatically generated

**Figure 2. Total Number of Cases and Number of decease by Age Quartile.** The bar chart shows the total number of cases (light grey) and the number of deceases (salmon) within each age quartile. Percentages and number of deceases are annotated above the respective bars.

A diagram of a data processing process

Description automatically generated

**Figure 3.** **Data Preparation Flowchart:** Steps involved in preparing data for machine learning. It includes importing necessary libraries, defining variables, splitting the data into training and testing sets, applying SMOTE to balance classes in the training set, and normalizing the data using MinMaxScaler. **Feature Selection & Modeling Flowchart:** flowchart outlines the process of feature selection and model training. It includes feature selection algorithms (RFE, SelectFromModel with various classifiers), training classifiers, creating a pipeline with feature selectors and classifiers using GridSearchCV, constructing a stacking model with base learners and a meta-classifier, and selecting the best feature selector and classifier.

A graph of a bar graph

Description automatically generated

**Figure 4. Correlation of Variables with Decease.** Horizontal bar chart illustrates the correlation values between various medical and procedural variables and the outcome of being deceased. Only variables with a correlation coefficient greater than 0.1 or less than -0.1 are included.

A collage of graphs and charts

Description automatically generated

**Figure 5.A** **Cross-Validation Scores:** cross-validation scores across five folds for the classifier ranging from approximately 0.975 to 0.998, indicating stable model performance across different subsets of the data. **B** **Model Evaluation:** Precision, recall, and F1-score for class 0 (non-deceased) are 0.94, 0.98, and 0.96 and for class 1 (deceased) are 0.91, 0.78, and 0.84, respectively. **C Receiver Operating Characteristic (ROC) Curve:** The AUC-ROC score is 0.93, indicating a high ability of the classifier to distinguish between the positive and negative classes. **D Confusion Matrix:** Heatmap displays the confusion matrix for the classifier LGBMClassifier with the feature selector SelectFromModel\_GB. True Negatives (700), False Positives (17), False Negatives (48) and True Positives (175). The overall accuracy of the model is 93%.

A screenshot of a graph

Description automatically generated

**Figure 6. Feature importances for the LGBMClassifier model**. The most influential features are EMERGENCY ADMISSION DATE TIME, AGE, and PRE-OPERATORY DAYS, with importance scores of 87, 82, and 54 respectively. Other notable features include EMERGENCY ADMISSION MONTH, DRG CODE, and BASE PROCEDURE ORDER DESCRIPTION.

A screenshot of a graph

Description automatically generated

**Figure 7.A** SHAP (SHapley Additive exPlanations) Summary Plot: impact of features on model predictions. Each point represents a Shapley value for a feature and an instance, with color indicating the feature value (red high, blue low). Features are ordered by the sum of SHAP value magnitudes across all samples, with the most influential features at the top. **B Individual Patient Case Analysis:**This force plot visualizes the impact of various features on the model's prediction for a specific patient (instance #988). The plot shows how each feature contributes to pushing the model's output from the base value (average prediction) to the final prediction for this individual case. This detailed breakdown helps interpret how different patient characteristics and clinical factors influence the model's decision for this particular case.

A screenshot of a graph

Description automatically generated

**Figure 8**. Comprehensive analysis of mortality rates across factors in surgical cases. **A** distribution of decease and total cases by pre-operative days. **B** different surgical procedures. **C** pre-operative days. **D** month of admission. Red bars represent deceased patients, grey portions show survivors, and labels indicate death counts and percentages. These visualizations provide insights into how factors such as procedure type, and admission timing and period correlate with mortality rates.