Honghui Hospital, Xi'an Jiaotong University

HoPreM Platform: Efficient Multimodal Multi-Task Prediction of Perioperative Events Following Hip Replacement Surgery

Study Protocol and Statistical Analysis Plan

Jingkun Liu
2024-11-1

Study Protocol and Statistical Analysis Plan

Study Protocol

Study Title:

Research on Multimodal Multi-Objective Integrated Machine Algorithm for Hip Replacement Surgery

Study Objectives:

The primary objective of this study is to develop a platform (HoPreM) that integrates multimodal data and employs multi-task learning to predict perioperative events in patients undergoing hip replacement surgery. These events include acute kidney injury (AKI), blood transfusion requirements, within 48-hour post-operative discharge (48hPOD), ICU transfer, and length of stay (LOS). The platform aims to enhance the prediction accuracy and improve clinical decision-making by leveraging various machine learning algorithms.

Study Design:

This is a retrospective cohort study involving data analysis from 4,722 patients who underwent hip replacement surgery between January 2020 and March 2024 at Xi'an Honghui Hospital. The study uses multimodal data, including demographic information, surgery-related data, medical history, and laboratory tests.

Inclusion Criteria:

- Patients aged 18 years or older
- Availability of preoperative and postoperative creatinine values
- Less than 10% missing values in medical records
- Logically consistent medical records

Exclusion Criteria:

- Patients younger than 18 years
- Non-hip replacement surgery patients
- Missing values greater than 10% in medical records
- Logical inconsistencies in medical records
- Absence of preoperative or postoperative creatinine values

Outcome Measures:

1. Primary Outcome Measures:
- Acute Kidney Injury (AKI) Incidence
- Blood Transfusion Requirements

- 48-Hour Postoperative Discharge (48hPOD)
- ICU Transfer
- Length of Stay (LOS)

2. Secondary Outcome Measures:
- The platform's predictive performance across these outcomes.

Data Collection:

Data will be collected from patient medical records, including demographic details, medical history, surgical data, and laboratory test results.

Statistical Analysis Plan

Data Preprocessing:

1. Missing Data Handling:
- For continuous numerical variables, missing data will be imputed using median values.
- For categorical variables, mode imputation will be used to ensure consistency.

2. Normalization:
- All continuous variables will be standardized using the StandardScaler to improve model convergence.

Model Development:

The multi-task prediction model will be constructed using the following methods:

1. Feature Selection:
- Lasso Regression and Random Forest will be used to select predictive features from multimodal data.
- Features selected include demographic data (e.g., age, weight), surgery-related data (e.g., surgical duration, blood loss), medical history (e.g., hypertension), and laboratory tests (e.g., creatinine, hemoglobin).

2. Ensemble Learning Models:
- The following ensemble learning algorithms will be compared for performance: XGBoost, Random Forest, LightGBM, and CatBoost.
- CatBoost will be selected as the optimal model based on its performance in multi-class classification and regression tasks.

Model Evaluation:

1. Classification Tasks:
- Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) will be used to evaluate the classification performance for AKI, blood transfusion, ICU transfer, and 48hPOD.

2. Regression Task (LOS):
- The regression performance will be evaluated using the Coefficient of Determination (R^2), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).

3. Precision-Recall (PR) Curves:
- PR-AUC will be calculated to evaluate model performance, especially in imbalanced data settings.

Feature Importance:

SHAP (Shapley Additive Explanations) values will be used to analyze the importance of features and their contribution to model predictions.

Statistical Analysis:

The Mantel test will be used to assess the associations between features and target variables.
The Pearson correlation matrix will be employed to analyze multicollinearity among features.

Validation and Testing:

The dataset will be split into training and testing sets with an 80:20 ratio to ensure generalizability.
K-fold cross-validation will be used to validate the model's robustness.
Early stopping will be employed to prevent overfitting during model training.

Software:

The analysis will be conducted using Python with libraries such as scikit-learn, XGBoost, CatBoost, LightGBM, and SHAP for feature selection, model training, and evaluation.