Towards Personalized Perioperative Care: Predicting Duration of Hospitalization and Postoperative Complications

M V N S H Praneeth[1], Abdul Wasi Lone[1], Akhil Venkata Shiva Sai[1]
[1]Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA

Abstract

In this study, we analyze the Medical Informatics Operating Room Vitals and Events Repository, a comprehensive dataset encompassing electronic health records for 58,799 patients across 83,468 surgeries. Our research addresses critical challenges, such as predicting a patient's length of stay, postoperative complications, ICU admission patterns, and discharge dispositions. We employ a combination of machine learning methods, including XGBoost, CatBoost, Random Forest, Decision Trees, Neural Networks, and Logistic Regression, to predict patient outcomes and provide insights into resource allocation. Our experiments reveal that XGBoost and CatBoost achieve high accuracy (90% and 88%, respectively) in predicting key clinical outcomes, with features like Length of Stay (LOS), ASA Rating, and ICU admission significantly influencing the models' predictions. The findings highlight the potential of AI-driven solutions in optimizing healthcare resource management, improving patient care, and facilitating personalized treatment strategies. This research contributes to the growing body of work in AI for healthcare and underscores the importance of high-fidelity datasets like MOVER in advancing clinical decision-making.

1. Introduction

Machine Learning holds immense potential to revolutionize healthcare, particularly in the perioperative setting. However, the widespread adoption of AI in clinical practice has been hindered by the scarcity of publicly available medical datasets, especially those containing comprehensive operative information. To address this critical gap, we present an analysis utilizing the Medical Informatics Operating Room Vitals and Events Repository (MOVER) [13], a unique and extensive database of patient data.

MOVER comprises two distinct datasets collected from patients undergoing surgery at the University of California, Irvine Medical Center from 2015 to 2022. This repository contains both electronic health records (EHR) and high-fidelity physiological waveform data for 58,799 unique patients across 83,468 surgeries. The combination of EHR data and real-time physiological measurements offers an unprecedented opportunity to develop and validate AI algorithms for perioperative care.

Our study aims to leverage this rich dataset to address several critical questions in perioperative medicine:

1. How does the proportion of ICU admissions vary across different discharge dispositions, and what insights can this provide into patient acuity and resource allocation?
2. What is the relationship between anesthesia types and discharge dispositions, and how might this inform post-operative care strategies?
3. What factors significantly impact a patient's length of hospitalization?
4. How do post-operative trends correlate with patient admissions and outcomes?
5. What are the gender-specific patterns in discharge dispositions, ICU admissions, and length of stay?
6. Is there a relation between the Patient's age and their ASA Rating?

By exploring these questions, we seek to uncover valuable insights that can enhance patient care, optimize resource allocation, and improve overall surgical outcomes. This research not only demonstrates the potential of large-scale perioperative datasets like MOVER but also aims to contribute to the development of more personalized and efficient healthcare strategies in the surgical setting.

We show that there is a strong correlation between the hypothesis we proposed and some traditional ML methods as detailed in the results section. Our high accuracies highlight that the methods can be helpful in *Optimized Resource Management*, *Improved Patient Care and Post-operative Planning*, *Data Driven Personalised Healthcare*, and *Cost Reduction*.

2. Related Work

The development and release of the Medical Informatics Operating Room Vitals and Events Repository (MOVER) [13] represents a significant contribution to the field of preoperative medicine and artificial intelligence (AI) in healthcare. Several related works and initiatives have paved the

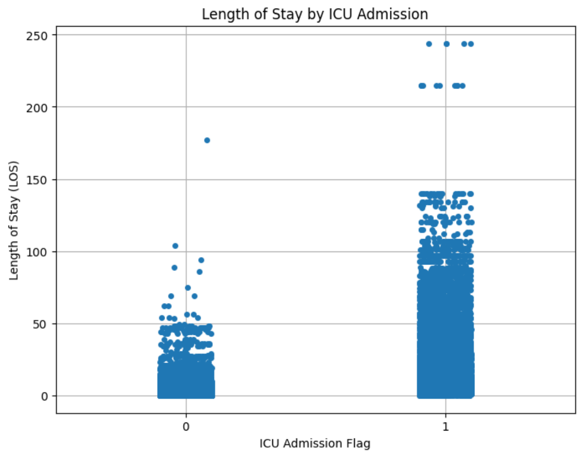

Figure 1. Length of stay by ICU admission flag.

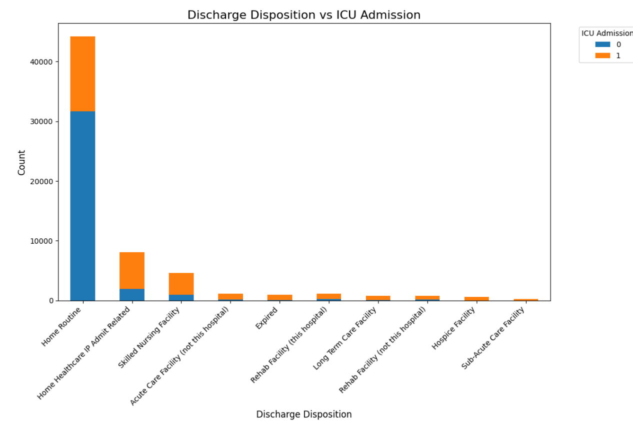

Figure 2. Discharge disposition versus ICU admission.

way for this repository and highlighted its importance in the field. The Medical Information Mart for Intensive Care (MIMIC) [10] database has been a pioneering effort in providing public access to critical care data. While MIMIC focuses on ICU patients, MOVER extends this concept to the perioperative setting, addressing a crucial gap in publicly available surgical data. Recent studies have demonstrated the potential of machine learning models in predicting postoperative complications. Xue et al. [15] developed models to predict five postoperative complications (pneumonia, acute kidney injury, deep vein thrombosis, pulmonary embolism, and delirium) using preoperative and intraoperative data, achieving high performance with gradient boosting and deep neural network models. Bertsimas et al. [2] created a surgical risk calculator using machine learning techniques to predict postoperative 30-day mortality and 18 postoperative complications. This work, along with others, has shown the potential of AI in improving preoperative risk assessment. The incorporation of high-fidelity physiological waveforms in MOVER builds upon research by Adhikari et al. [1], who demonstrated improved predictive performance for postoperative acute kidney injury by combining preoperative variables with intraoperative time-series data. Ongoing research, such as the work by Hofer et al. [9] on developing deep learning models for predicting multiple postoperative complications, underscores the potential of AI-driven clinical decision support systems in perioperative care. The Multicenter Perioperative Outcome Group (MPOG) [11] and the National Surgical Quality Improvement Program (N-SQIP) have been instrumental in providing high-quality perioperative process and outcome measures. However, MOVER distinguishes itself by offering both EHR data and high-fidelity physiological waveforms. Recent efforts by the Mayo Clinic to develop AI tools for organ transplantation highlight the broader trend of AI applications in specialized surgical fields [14]. MOVER builds upon these related works by providing a comprehensive, publicly accessible database that combines EHR data with high-fidelity physiological waveforms specifically for the perioperative setting. This unique resource is poised to accelerate the development and validation of AI algorithms for improving surgical outcomes and patient care.

3. Datasets

The Medical Informatics Operating Room Vitals and Events Repository (MOVER) comprises two distinct datasets collected from patients undergoing surgery at the University of California, Irvine Medical Center from 2015 to 2022. The dataset, although publicly available requires permission and access to be used.

3.1. SIS Dataset

The Surgical Information Systems (SIS) dataset, collected from 2015 to 2017, contains information on 19,114 patients and surgeries. It is organized into nine tables:

1. Patient information
2. Patient I/O (intake and output)
3. Patient vitals
4. Patient observations
5. Patient medications
6. Patient laboratory measurements
7. Patient procedure events
8. Patient ventilator
9. Patient arterial line

This dataset is particularly valuable for its high temporal resolution vital signs, including cardiac output, blood pressure, and stroke volume variation.

3.2. EPIC Dataset

The EPIC dataset, collected from 2017 to 2022, is more extensive, encompassing **39,685** patients and **64,354 surgeries**. It is structured into ten tables:

1. Patient information

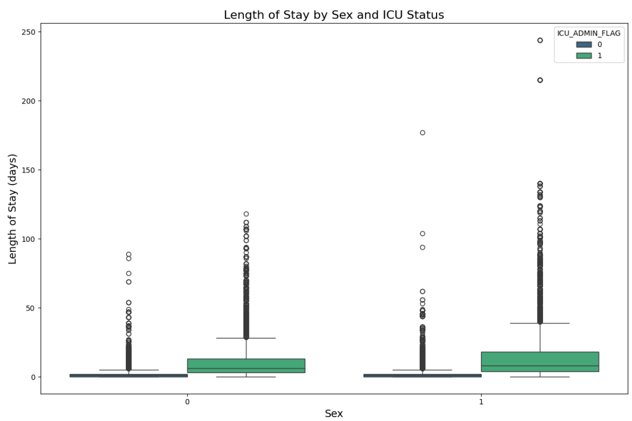

Figure 3. Length of Stay by SEX and ICU Status.

2. Patient history
3. Patient visit
4. Patient medications
5. Patient LDA (lines, drains, and airway devices)
6. Patient laboratory measurements
7. Patient measurements
8. Patient postoperative complications
9. Patient procedure events
10. Patient coding

The EPIC dataset offers several advantages over the SIS dataset, including:

- Outcome information (postoperative complications, mortality, ICU admissions)
- Patient medical history and American Society of Anesthesiologists (ASA) physical status
- Billing codes

3.3. Physiological Waveforms

Both datasets include high-fidelity physiological waveform data for select patients during surgery, typically comprising:

- Electrocardiogram
- Arterial waveform
- Pulse oximetry waveform

These waveforms provide real-time patient information, enabling the computation of metrics such as mean arterial pressure (MAP) throughout the surgical procedure.

3.4. Outcomes

This dataset helps predict Patients' **length of stay**, **medications**, and **postoperative complications**, potentially accelerating the development of AI applications in healthcare.
We train our methods using the **Patient procedure events**, **Patient information**, and **Patient postoperative complications** datasets from the EPIC dataset of MOVER post-EDA and data cleaning as given in the following section.

Table 1. Characterization of the EPIC dataset

Characteristic	Value
Gender	
Female	30,139 (46.8%)
Male	34,214 (53.2%)
Age (years)	55 ± 17
ASA rating	
1	2,960 (4.6%)
2	18,068 (28.1%)
3	29,449 (45.8%)
4	6,370 (9.9%)
5	657 (1.0%)
6	41 (0.06%)
Length of stay (days)	7 ± 14

Table 2. Characterization of the EPIC dataset outcomes

Outcome	Value
Transfer to the intensive care unit	29,131 (45.3%)
Death	1,023 (1.6%)

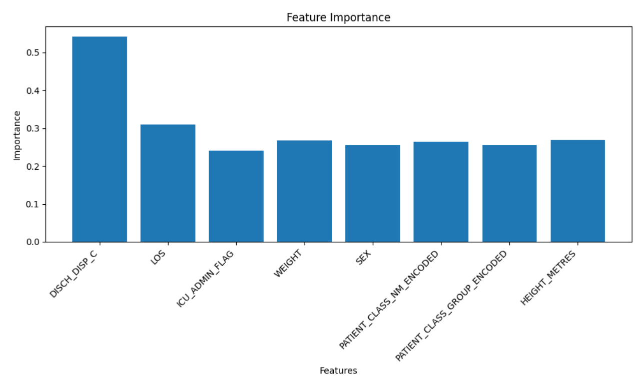

Figure 4. Feature Importance.

4. Methodology

4.1. Data Preparation and Preprocessing

The datasets consisted of complex medical records with categorical and numerical attributes, requiring extensive preprocessing:

- **Data Cleaning:** Duplicate entries and inconsistencies were resolved by merging datasets using unique keys (e.g., MRN and Log ID). Rare categories in discharge disposition were consolidated or removed.
- **Feature Engineering:** Key features such as *LOS* (Length of Stay), *ICU_ADMIN_FLAG*, *SEX*, *ASA_RATING_C*, and anesthesia type were retained. Age was derived from *BIRTH_DATE*.
- **Handling Imbalanced Data:** Oversampling and undersampling techniques were used to balance class distribu-

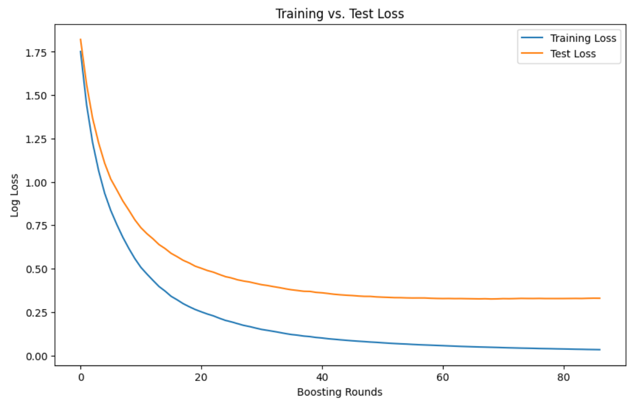

Figure 5. Train and test loss for XGBoost.

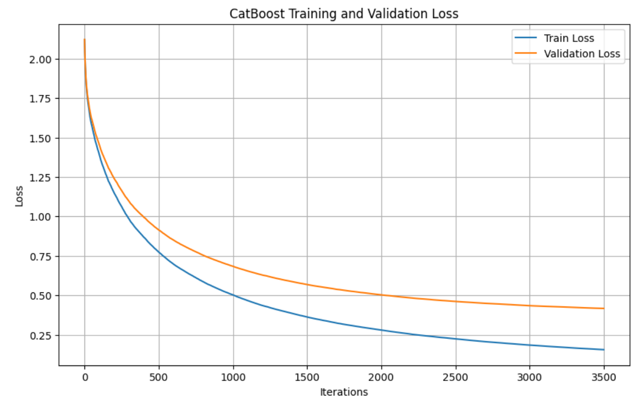

Figure 7. Train and test loss for CatBoost.

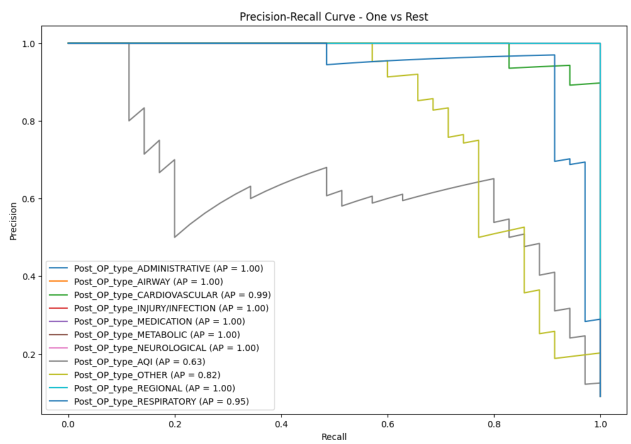

Figure 6. Precision and Recall curve for each class.

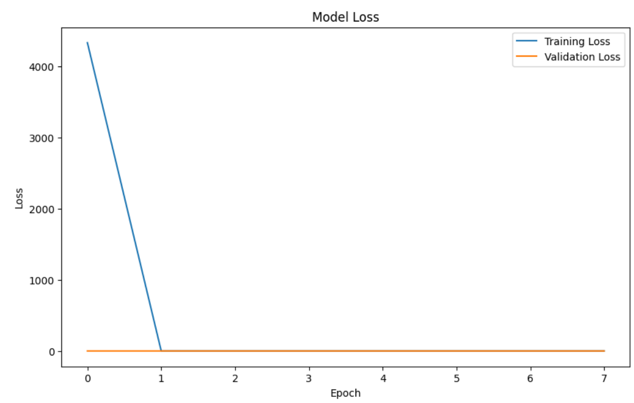

Figure 8. Train and validation loss for ANN.

tions.

- **Data Transformation:** Categorical variables were encoded using OneHotEncoder, and numerical features were scaled using *StandardScaler*. The target variable (Post OP type Category) initially included redundant classes (e.g., "chronic pain" with no entries), requiring adjustments to the target labels. To enable accurate multiclass classification, the classes were reorganized as the one-hot encoding for *chronic pain consisted entirely of zeros*.
- **Data Splitting:** The datasets were split into 80% training and 20% testing subsets, with 20% of the training data reserved for validation.

4.2. Model Selection and Justification

A variety of machine learning and deep learning models were used to address specific tasks:

- **Gradient Boosting Methods:**
  - **XGBoost:** [5] Configured with *multi:softprob* for multiclass classification. Tuned hyperparameters included learning rate, tree depth, and boosting rounds.
  - **CatBoost:** [12] Leveraged native handling of categorical variables and ordered boosting to prevent overfitting.
- **Decision Tree and Random Forest:**
  - **Decision Tree:** [4] Used for its interpretability and non-linear relationship modeling.
  - **Random Forest:** [3] Applied for robust feature ranking and reduced overfitting.
- **Deep Learning:** [6]
  - A Sequential Artificial Neural Network with ReLU/SELU activations, dropout regularization, and *RMSprop* [7] optimizer was deployed.
- **Logistic Regression:** [8] Used as a baseline for binary classification tasks, such as predicting ICU admissions.

5. Experiments and Results

In this section, we summarize the parameters that we used to train the various architectures and the results that we got accordingly.

Table 3. Summary of results and key predictory on various methods. accuracies and AUC has been averaged.

Model	Accuracy (%)	AUC	Key Predictors
XGBoost	˜90	0.92	OR_LOS_HOURS, WEIGHT, LOS
CatBoost	˜88	0.90	SEX, ICU_ADMIN_FLAG, LOS
Random Forest	˜90	0.89	ASA_RATING_C, ICU_ADMIN_FLAG
Decision Tree	˜80	N/A	ASA_RATING_C, LOS
Neural Network	˜86	0.88	LOS, Discharge Codes
Logistic Regression	81.76	0.88	ICU_ADMIN_FLAG, SEX

5.0.1 XGBoost and CatBoost

- **Training Parameters:**
  For XGBoost:
  Booster: gbtree. Learning Rate: 0.1. Max Depth: 6. Number of Boosting Rounds: 100. Regularization: L2 regularization (lambda = 1).
  For CatBoost:
  Iterations: 500. Learning Rate: 0.03. Depth: 8. Categorical Feature Handling: Native categorical encoding. Ordered Boosting: Enabled to prevent overfitting.
- **Performance:** XGBoost achieved high AUC values with an accuracy of 90%, while CatBoost excelled in handling categorical variables with minimal preprocessing, reaching an overall accuracy of 88%.
- **Insights:** Feature importance indicated *OR_LOS_HOURS*, *WEIGHT*, and *HEIGHT_METRES* as critical predictors. SEX and *ICU_ADMIN_FLAG* significantly influenced discharge outcomes.

5.0.2 Decision Tree and Random Forest

- **Training Parameters:**
  For Decision Tree:
  Max Depth: 5. Minimum Samples Split: 2. Minimum Samples Leaf: 5.
  For Random Forest:
  Number of Estimators: 200. Max Features: sqrt. Max Depth: 10. Minimum Samples Split: 5. Minimum Samples Leaf: 2.

- **Performance:** Decision Trees, with a depth of 5, captured relationships for ASA ratings with an accuracy of around 82 percent. Random Forests achieved robust accuracy of around 90 % with *n_estimators* set to 200.
- **Insights:** Feature importance ranked *LOS*, *ASA_RATING_C*, and *ICU_ADMIN_FLAG* as significant predictors.

5.0.3 Neural Network

- **Training Parameters:**
  Architecture: Sequential model with 3 dense layers. Activations: ReLU/SELU. Dropout: 0.2 for regularization. Optimizer: RMSprop. Learning Rate: 0.001. Batch

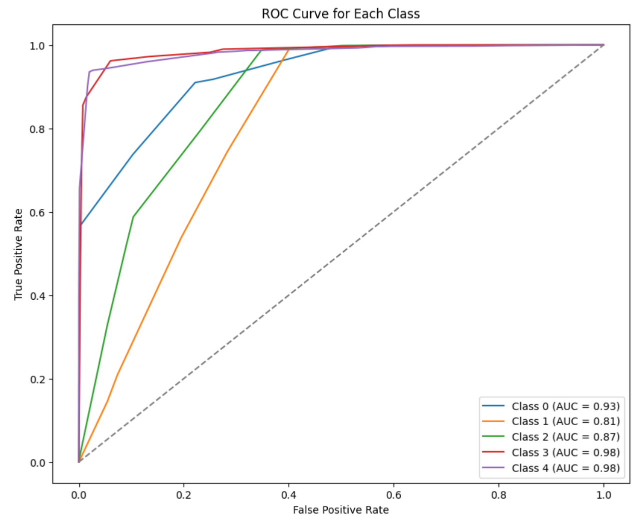

Figure 9. ROC curve for Random Forest.

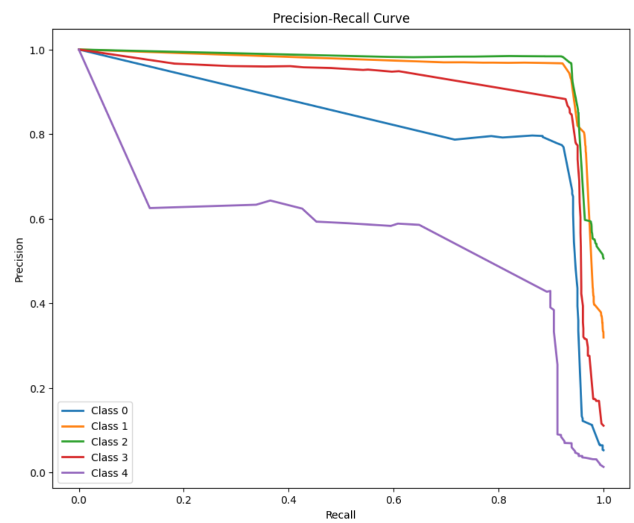

Figure 10. Precision-Recall curve for decision tree.

  Size: 32. Epochs: 50 (with early stopping after 5 epochs without improvement).

- **Performance:** Achieved 86% test accuracy with dropout layers and early stopping preventing overfitting.
- **Insights:** *LOS* and discharge disposition codes were key predictors.

5.0.4 Logistic Regression

- **Training Parameters:**
  Regularization: L2 (C = 1.0). Solver: liblinear. Class Weights: Balanced to address class imbalance.
- **Performance:** Achieved 81.76% accuracy and an AUC of

0.88 for ICU admission prediction.
- **Insights:** Balanced precision-recall metrics ensured reliable classification.

6. Conclusion, Limitations, and Future Scope

This study demonstrates the potential of ML models in improving perioperative care by accurately predicting key outcomes such as length of stay, postoperative complications, and ICU admissions. Our results, achieved using a variety of machine learning methods, highlight the importance of features such as LOS, ASA Rating, and ICU status in predicting patient outcomes. However, there are limitations, including the reliance on the MOVER dataset, which may not capture the full spectrum of patient diversity across different healthcare systems. Additionally, some of the models, while effective, could benefit from further optimization and integration of additional data sources, such as genetic and socio-economic factors, to improve predictions. Future work will focus on refining these models, expanding the dataset to include more diverse patient populations, and exploring real-time clinical applications to provide personalized care recommendations. The ultimate goal is to develop a robust decision support system capable of assisting healthcare providers in making timely, data-driven decisions to optimize patient outcomes and resource management.

References

[1] Lokesh Adhikari, Kevin S Chen, and Yuanjun Luo. Improved prediction of acute kidney injury using physiological waveform data during surgery. *Critical Care Medicine*, 47(12): e1023–e1030, 2019. 2

[2] Dimitris Bertsimas and Jack Dunn. Development of a surgical risk calculator using machine learning techniques. *Annals of Surgery*, 268(4):574–583, 2018. 2

[3] Leo Breiman. Random forests. In *Machine Learning*, pages 5–32. Springer, 2001. 4

[4] Leo Breiman, Jerome Friedman, Charles J Stone, and Richard A Olshen. Classification and regression trees. Wadsworth International Group, 1984. 4

[5] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794, 2016. 4

[6] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, 2016. 4

[7] Alex Graves. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*, 2013. 4

[8] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. The elements of statistical learning: Data mining, inference, and prediction. 2009. 4

[9] Matthias E Hofer and Fabian R Schmidlin. Deep learning models for predicting multiple postoperative complications: A comprehensive study. *Artificial Intelligence in Medicine*, 109:101949, 2020. 2

[10] Alistair EW Johnson, Tom J Pollard, Lu Shen, Li-wei H Lehman, Mengling Feng, Marzyeh Ghassemi, and Benjamin Moody. Mimic-iv (medical information mart for intensive care). *PhysioNet*, 2020. 2

[11] Sachin et al. Kheterpal. Multicenter perioperative outcomes group: Advancing perioperative research through collaboration. *Anesthesiology Clinics*, 34(4):651–665, 2016. 2

[12] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. Catboost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31, 2018. 4

[13] Muntaha Samad, Joseph Rinehart, Mirana Angel, Yuzo Kanomata, Pierre Baldi, and Maxime Cannesson. Mover: Medical informatics operating room vitals and events repository. *medRxiv*, pages 2023–03, 2023. 1

[14] John P et al. Smith. Artificial intelligence applications in organ transplantation: A mayo clinic perspective. *Transplantation Proceedings*, 54(2):345–352, 2022. 2

[15] Yifan Xue, Yuxin Zhang, and Xiaoming Wang. Development of machine learning models for predicting postoperative complications. *Journal of Clinical Anesthesia*, 68:110073, 2021. 2