



# Predictive Modeling of Anesthesia Types and Their Associations with Clinical Outcomes Using the MIMIC-IV Database

Goli Asgari<sup>1</sup>, Ata Amini<sup>2</sup>, Pooya Vahedi<sup>3</sup>, Yasmin Tabatabavakili<sup>2</sup>, Jan Muzik<sup>4,5</sup>, Elham Nazari<sup>6</sup>,  
\*, Tahmineh Aldaghi<sup>4</sup>,\*\*

<sup>1</sup> Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>2</sup> Department of Health Information Management and Medical Informatics, School of Allied Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran

<sup>3</sup> School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>4</sup> Institute of Biophysics and Informatics, First Faculty of Medicine, Charles University, Prague, Czech Republic

<sup>5</sup> Department of Information and Communication Technologies in Medicine, Faculty of Biomedical Engineering, Czech Technical University, Prague, Czech Republic

<sup>6</sup> Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

\*Corresponding Author: Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: elham.nazari@sbmu.ac.ir

\*\*Corresponding Author: Institute of biophysics and informatics, First faculty of medicine, Charles University, Prague, Czech Republic. Email: tahmineh1989@gmail.com

Received: 13 September, 2025; Revised: 25 October, 2025; Accepted: 31 October, 2025

## Abstract

**Background:** Perioperative management plays a pivotal role in reducing complications and improving patient outcomes. The choice of anesthesia can significantly influence postoperative trajectories; however, current clinical practice lacks data-driven tools to support personalized selection. With the emergence of machine learning (ML) and access to large clinical databases, new opportunities have arisen to optimize anesthesia selection and predict patient outcomes.

**Methods:** This retrospective study utilized the MIMIC-IV database (version 3.1), which comprises data from over 364,000 patients. After feature selection and preprocessing, 28 demographic, clinical, and laboratory variables were used to model four types of anesthesia (general, regional, local, and sedation) in 31,821 patients. Supervised models [random forest, logistic regression, support vector machine (SVM), k-nearest neighbors (KNN), decision tree, gradient boosting, and extreme gradient boosting (XGBoost)] were trained to predict anesthesia type; synthetic minority oversampling technique (SMOTE) was applied within training to address class imbalance. Performance was evaluated on a held-out test set using accuracy and macro-F1. Associations between anesthesia type and short-term outcomes (in-hospital mortality, 30-day readmission, postoperative infection, and length of stay) were examined through observational comparisons.

**Results:** Boosting-based models achieved the highest predictive performance, with XGBoost and gradient boosting reaching approximately 83% accuracy on the test set (macro-F1 for XGBoost = 0.45). Across anesthesia groups, regional anesthesia was associated with more favorable outcomes than general anesthesia – namely lower in-hospital mortality, fewer infections, shorter hospital and intensive care unit (ICU) stays, and reduced 30-day readmission – although estimates for the regional group were less precise due to small sample size and residual confounding is possible.

**Conclusions:** Anesthesia type was associated with short-term clinical outcomes in this large, single-center cohort, and boosting models demonstrated strong predictive performance for anesthesia type. These findings suggest a potential role for predictive analytics in perioperative planning, while underscoring that observed differences are associations rather than causal effects. Prospective multicenter studies and randomized trials, with longer-term patient-reported outcomes, are warranted. Predictive models require external validation and calibration before they can be used clinically.

**Keywords:** XGBoost Model, MIMIC-IV Database, Non-invasive Anesthesia Techniques

## 1. Background

The management of the perioperative period represents a critical component of healthcare, involving the coordinated care of patients across the preoperative, intraoperative, and postoperative phases. Well-structured management in this setting can reduce

complications, decrease length of stay, and lower the risks of readmission and mortality (1, 2). Anesthesia technique plays a pivotal role in postoperative outcomes, including pain, nausea and vomiting, cognitive recovery, respiratory function, and length of stay. These factors guide clinicians in selecting protocols that promote optimal recovery (3). Recent

advancements, including target-controlled infusion, ultrasound-guided regional anesthesia, and closed-loop anesthesia systems, have improved anesthetic safety and efficacy. However, many of these innovations are still under development and not widely adopted in routine practice (4, 5). Incorporating these innovations into practice, supported by sustained research and technological adoption, is essential to further improving patient outcomes (6).

As healthcare systems become increasingly digitized, large clinical datasets have emerged, enabling new approaches to perioperative research. Over the past decade, artificial intelligence (AI) and machine learning (ML) have garnered increasing interest in the medical field. In perioperative care, ML can generate predictive models and early warning systems, support the identification of critical illness, and improve the management of high-risk patients (7-10). Unlike traditional statistical methods, AI can learn from data, refine its performance over time, and develop models tailored to individual patients (11). The intensive care unit (ICU), due to continuous patient monitoring, generates large amounts of clinical data. Publicly available datasets, particularly the MIMIC series, have enabled extensive research in critical care. MIMIC-I included over 90 patients (12), followed by MIMIC-II, which featured larger cohorts and digital integration (13), and MIMIC-III, which expanded to more than 40,000 patients (14). The MIMIC-IV dataset, which covers ICU admissions from 2008 to 2019, provides precise digital information, including electronic medication records, and features a modular structure that facilitates efficient use in clinical research (15).

Despite extensive interest in anesthesia-outcome relationships, most prior studies have been limited to specific procedures or small cohorts, and few have simultaneously assessed both predictive feasibility and real-world outcome associations on a large scale.

## 2. Objectives

This study aims to evaluate whether routinely collected perioperative clinical features can accurately predict anesthesia type and to investigate the association between anesthesia modality and short-term clinical outcomes using the MIMIC-IV dataset and modern ML techniques.

## 3. Methods

### 3.1. MIMIC-IV Database

This study utilized version 3.1 of the MIMIC-IV database, a comprehensive, structured, and anonymized dataset created by the MIT Laboratory for Computational Physiology and made freely available to researchers via the PhysioNet platform. All data used in this study are publicly available to credentialed researchers through the PhysioNet platform. The database contains extensive information about patients who were treated in the emergency room and ICU at Beth Israel Deaconess Medical Center in Boston, Massachusetts. Version 3.1 has information on 364,627 unique patients, 546,028 hospital admissions, and 94,458 ICU stays. It is divided into two main modules: HOSP, which has hospital-level data like admissions, diagnoses, laboratory tests, medications, and administrative or billing information; and ICU, which has continuous ICU data like vital signs, clinical events, administered medications, and caregiver interventions. Only researchers who have finished CITI training on "Data or Specimens Only Research" and signed the data use agreement can access this database. The MIMIC-IV database is a publicly available, de-identified dataset. Therefore, this study did not require separate approval from an Institutional Review Board. MIMIC-IV is now one of the most important tools for research in ML, clinical risk analysis, outcome prediction, and improving medical processes (15).

### 3.2. Collecting Data and Choosing Features

After signing up and obtaining permission on the PhysioNet platform, data analysis was performed in the Jupyter Notebook environment using Python version 3.11. The libraries used were Pandas, NumPy, and scikit-learn. The unique identifier ( $\text{subject}_{id}$ ) was used to get patient-level information. Initially, 32 candidate features were selected based on their clinical significance and recommendations from anesthesiology experts. We used a combination of domain expertise and statistical filtering to select the features. Initially, 32 candidate features were chosen based on clinical relevance, followed by statistical analysis using ANOVA. These included demographic variables (like sex and anchor age), information about the admission (like the type of

admission, insurance, in-hospital death rate, and 30-day readmission rate), comorbidities (like diabetes, hypertension, CAD, cancer, CKD, COPD, sepsis, and infections, which were coded as binary variables), and procedural information (like the type of surgery and anesthesia). We averaged laboratory and physiological measures, including hemoglobin, platelet count, creatinine, blood urea nitrogen (BUN), sodium, potassium, INR, diastolic blood pressure, GCS score, and weight, to create patient-level summaries that were representative. Records containing incomplete or implausible values were omitted. After that, a univariate statistical analysis using ANOVA was done, and variables with  $P < 0.05$  were thought to be statistically significant. A total of 28 final features were chosen for modeling.

### 3.3. Target Variable and Outcomes

The main variable of this study was the type of anesthesia, which was a multiclass categorical variable with four groups: General, regional, local, and sedation. Furthermore, a series of secondary outcomes was included to facilitate a more thorough assessment of anesthesia-related effects. These outcomes consisted of duration of hospital stay, duration of ICU stay, in-hospital mortality, 30-day readmission, and incidence of infection. The lengths of hospital and ICU stays were broken down into three groups: Short ( $\leq 24$  hours), intermediate (24 - 72 hours), and long ( $> 72$  hours). In-hospital mortality and 30-day readmission were classified as binary outcomes, and postoperative or in-hospital infection was similarly regarded as a binary variable. This multidimensional outcome design facilitated a more comprehensive examination of the correlation between anesthesia type and clinical outcomes.

### 3.4. Data Preprocessing

The LabelEncoder converted categorical variables, such as sex, admission type, and surgery type, into numerical values, enabling the data to be used for modeling. We used MinMaxScaler to normalize the numerical variables so that differences in scale would have less of an effect. Data normalization was performed using MinMaxScaler for consistency across models. However, for tree-based models like extreme gradient boosting (XGBoost) and random forest, this scaling does not affect model performance; it was applied to ensure

compatibility with other models. We averaged temporal data, such as lab results and vital signs, to make things easier and more consistent. Using column means, SimpleImputer filled in missing numbers, and the Z-score method identified and removed outliers. The last dataset was divided into two parts: A training set and a testing set, with an 80/20 split. A fixed random state was used to ensure the results could be reproduced. To fix the problem of class imbalance in anesthesia type, two methods were used: The synthetic minority oversampling technique (SMOTE) to add more examples to minority classes and adjusting class weights ( $\text{class\_weight} = \text{'balanced'}$ ) for some models to make them more sensitive to categories that are not well represented.

### 3.5. Model Implementation

All modeling procedures were executed in scikit-learn using Jupyter Notebook. The primary goal was to determine the type of anesthesia to use based on demographic, clinical, and surgical factors, and to examine the relationship between these factors and secondary outcomes. The models used were random forest, logistic regression, support vector machine (SVM), k-nearest neighbors (KNN), decision tree, gradient boosting, and XGBoost, among others. Multi-output models, such as MultiOutputClassifier and MultiOutputRegressor, were used when multiple outcomes needed to be predicted simultaneously.

### 3.6. Model Evaluation

We evaluated the model on a test set to assess its performance. Standard multiclass evaluation metrics, such as overall accuracy, precision, recall, and F1 score, were used to predict the type of anesthesia as a four-class categorical variable. We utilized functions from the scikit-learn library to perform all the mathematical operations. Lastly, the models that performed best in predicting the type of anesthesia were selected as the best-performing models.

## 4. Results

### 4.1. Baseline Characteristics

After data preprocessing and outlier removal, a total of 31,821 patients were included in the study. These

**Table 1.** Baseline Characteristics <sup>a</sup>

Features	General Anesthesia (N = 24,545)	Regional Anesthesia (N = 127)	Local Anesthesia (N = 2990)	Sedation (N = 4159)
Age	62.4 ± 14.8	65.7 ± 13.1	62.1 ± 15.0	61.7 ± 13.5
Gender (male)	52.1	53.0	54.5	50.5
Diabetes	23.4	25.6	20.2	21.8
Hypertension	39.8	47.3	42.2	39.3
Coronary artery disease	15.2	16.1	14.3	12.8
Kidney failure	12.5	13.3	11.1	10.0
Asthma and COPD	10.8	13.2	9.5	8.7
Cancer	6.2	7.8	5.9	6.0
Sepsis history	7.1	8.3	6.7	6.1
Hemoglobin (g/dL)	13.2 ± 1.8	13.5 ± 1.9	13.3 ± 1.7	13.0 ± 1.8
Platelets ( $\times 10^3/\mu\text{L}$ )	250 ± 35	270 ± 40	245 ± 30	235 ± 38
Creatinine (mg/dL)	1.1 ± 0.3	1.0 ± 0.2	1.1 ± 0.3	1.2 ± 0.3
Infection history	9.1	4.6	6.1	5.5
Hospital length of stay (d)	2.6 ± 4.9	1.8 ± 3.7	2.4 ± 4.2	2.3 ± 3.5
ICU length of stay (d)	2.1 ± 4.2	1.4 ± 3.5	2.0 ± 3.7	2.2 ± 3.1

Abbreviation: ICU, intensive care unit.

<sup>a</sup> Values are expressed as mean ± SD or percentage.

patients were divided into four groups based on the type of anesthesia. The general anesthesia group had the largest number of patients (24,545), followed by the sedation, local, and regional anesthesia groups, which ranked second, third, and fourth, with 4,159, 2,990, and 127 patients, respectively. The mean age of the patients was approximately 62.4 years. In terms of gender, 52.1% of the subjects were male and 47.9% were female.

A significant difference was observed between the anesthesia groups. Patients in the regional anesthesia group were older on average than those in the general anesthesia group (mean age 65.7 vs. 62.4;  $P < 0.001$ ). Also, the prevalence of hypertension was higher in the regional anesthesia group (47.3% vs. 39.8%;  $P < 0.001$ ), while no significant difference was observed for diabetes ( $P = 0.21$ ). The prevalence of CAD differed modestly but significantly across groups ( $P = 0.008$ ). Most laboratory parameters showed statistically significant between-group differences, although absolute differences were small.

Table 1 shows the baseline characteristics (demographic, clinical, and laboratory) by type of anesthesia performed. This table provides information on the mean age, gender, comorbidities (including diabetes, hypertension, coronary artery disease, renal failure, and cancer), and laboratory parameters (hemoglobin, platelets, creatinine, and infection) for

each anesthesia group. Significant differences in demographic and clinical characteristics can affect clinical outcomes.

#### 4.2. Performance of Machine Learning Models (Synthetic Minority Oversampling Technique-Balanced Data)

Using 28 selected features, seven ML models were compared on the held-out test set after applying SMOTE to address class imbalance across anesthesia types. Gradient boosting and XGBoost achieved the highest overall accuracy (83.1% and 83.1%, respectively), with random forest close behind (82.6%). At the class level, gradient boosting yielded the best F1 score for general anesthesia, whereas XGBoost outperformed the others for sedation and local anesthesia; all models struggled with regional anesthesia due to the extremely small sample size. Overall, KNN and decision tree underperformed relative to boosting-based models.

#### 4.3. Data Balance (Synthetic Minority Oversampling Technique) and Class Weights

The SMOTE increased overall accuracy (from approximately 77.5% without resampling to approximately 83% with resampling for boosting-based models), but did not significantly improve macro-F1, and the performance for the regional class remained low. We therefore report the SMOTE results as the main

**Table 2.** Performance of Machine-Learning Models on the Synthetic Minority Oversampling Technique-Balanced Test Set

Models	Accuracy	Macro-Precision	Macro-Recall	Macro-F1	Weighted-F1
Gradient boosting	0.8314	0.52	0.42	0.43	0.80
<b>XGBoost</b>	<b>0.8308</b>	<b>0.60</b>	<b>0.43</b>	<b>0.45</b>	<b>0.80</b>
Random forest	0.8259	0.52	0.38	0.38	0.78
<b>KNN</b>	<b>0.7832</b>	<b>0.40</b>	<b>0.33</b>	<b>0.34</b>	<b>0.74</b>
Decision tree	0.7361	0.38	0.39	0.38	0.74
<b>SVM</b>	<b>0.6573</b>	<b>0.40</b>	<b>0.53</b>	<b>0.41</b>	<b>0.71</b>
Logistic regression	0.5527	0.37	0.52	0.35	0.63

Abbreviations: XGBoost, extreme gradient boosting; KNN, k-nearest neighbors; SVM, support vector machine.

analysis and provide non-resampled results as a sensitivity analysis. Where supported (e.g., LR, SVM, and RF),  $\text{class}_{\text{weight}} = \text{'balanced'}$  showed broadly consistent patterns with the SMOTE analysis.

#### 4.4. Main Result

The XGBoost and gradient boosting achieved the highest performance in predicting the type of anesthesia. For XGBoost, the approximate binomial 95% CI for accuracy was 82.2 - 84.0%; for gradient boosting, 82.2 - 84.1%. The macro-averaged F1 score for XGBoost was 0.45, reflecting the impact of severe class imbalance. Given the near-tie in accuracies, we report both boosting models as top performers and provide full per-class metrics in Table 2. Agreement between predictions and true labels, measured by Cohen's  $\kappa$ , was 0.48 (95% CI 0.46 - 0.51) for XGBoost and 0.47 (0.45 - 0.49) for gradient boosting on the four-class task.

#### 4.5. Clinical Interpretation

This study indicates that the choice of anesthesia is associated with differences in clinical outcomes, such as in-hospital mortality, 30-day readmission, infection, and length of hospital and ICU stay. Regional anesthesia was associated with more favorable outcomes compared with general anesthesia. In our analyses, regional anesthesia was associated with approximately 37% lower odds of in-hospital mortality ( $\text{OR} \approx 0.63$ ), shorter ICU stays, fewer infections, and lower 30-day readmission and hospital length of stay. Possible mechanisms — such as attenuation of systemic inflammatory and stress responses — are biologically plausible but remain hypotheses. Because general anesthesia is often selected for more complex or urgent procedures, residual confounding (e.g., confounding by indication) may

persist, and estimates for the regional group are limited by its small sample size. Therefore, these findings should be interpreted as associations rather than causal effects, while still providing information that may help clinicians select the most appropriate anesthetic technique for patients with specific clinical conditions.

## 5. Discussion

In this large retrospective cohort study, which included more than 30,000 patients from the MIMIC-IV database, we observed that the type of anesthesia was associated with short-term outcomes. Compared with general anesthesia, regional anesthesia was associated with lower in-hospital mortality, decreased infection rates, shorter hospital and ICU stays, and reduced 30-day readmissions. Furthermore, ML analyses demonstrated that demographic and clinical characteristics could predict anesthesia type with high accuracy. The XGBoost and gradient boosting achieved the highest accuracy (approximately 83%) on the held-out test set (macro-F1 for XGBoost  $\approx 0.45$ ). Collectively, these findings highlight both the clinical implications and the predictive utility of anesthesia selection in patient stratification.

It is important to emphasize that this study is observational in design, and our primary goal was to identify patterns and associations rather than establish causality. Our findings are consistent with earlier evidence that regional anesthesia is associated with more favorable perioperative outcomes than general anesthesia, although effect sizes may vary across subgroups and outcomes. Earlier studies, including systematic reviews and meta-analyses, have compared the efficacy and patient outcomes associated with general, regional, and local anesthesia in both surgical and critical care settings. For example, recent meta-



analyses found no significant difference in overall sedation or anesthesia success rates between remimazolam and propofol (16). However, remimazolam may reduce the risk of hypoxemia and injection pain, at the expense of longer awakening times (17).

Regional anesthesia techniques, such as peripheral nerve blocks, have been associated with lower early postoperative pain scores compared with general anesthesia, though differences diminish after the first 12 hours post-surgery. Moreover, regional approaches are associated with lower opioid consumption and fewer instances of nausea and vomiting immediately postoperatively (18). When comparing regional and general anesthesia for major procedures, Bayesian meta-analyses suggest that the use of dexmedetomidine as an adjunct can further improve quality of recovery (QoR) and patient-centered outcomes, including reduced agitation and faster return to baseline function (19).

More recently, analyses using large critical care databases, such as MIMIC-IV, have produced complementary associations regarding anesthetic techniques and sedative combinations. These studies demonstrate that sedatives, such as dexmedetomidine, have been associated with improved survival and fewer complications compared to midazolam or propofol in mechanically ventilated ICU patients (16, 20, 21). Other studies using the same dataset have also suggested that ketamine may offer short-term mortality benefits in critically ill patients on vasopressors, though its advantages may not persist at 90 days (22).

Combination anesthesia approaches, such as general anesthesia combined with regional blocks, have been associated with enhanced recovery markers compared to general anesthesia alone, including improved postoperative pulmonary function, more stable hemodynamics, lower complication rates, and faster recovery of cognitive function and sleep quality. For instance, patients receiving combined anesthesia exhibited better pulmonary oxygenation, more stable hemodynamics, and a faster recovery of cognitive and sleep quality after surgery compared to those receiving general anesthesia alone. Additionally, combined general and epidural anesthesia improved pain control and psychomotor recovery after major surgery (23-25).

Additionally, new agents, such as remimazolam, are being investigated for their role in reducing

postoperative nausea and vomiting, which remain important recovery outcomes in anesthesia practice (26). Because most of these analyses are observational and often conducted in single centers, residual confounding and selection bias remain possible; therefore, effect sizes should be interpreted cautiously.

What distinguishes the present study from earlier literature is its scale, breadth, and methodological approach. Unlike prior single-procedure or limited-population analyses, our work leveraged data from more than 30,000 patients undergoing diverse surgical interventions. By integrating modern ML approaches, we not only documented between-group outcome differences across anesthesia modalities but also demonstrated that patient- and procedure-level factors can predict anesthesia type with high accuracy (XGBoost and gradient boosting achieved approximately 83% accuracy on the held-out test set). These findings contribute to the growing body of evidence suggesting that anesthesia choice is not merely a procedural decision but is also associated with the patient's trajectory, and they highlight the potential role of predictive modeling in guiding perioperative care.

In contrast to randomized controlled trials that focus on narrow patient groups, our database-driven approach provides insights into real-world practice patterns and outcomes across a wide range of clinical contexts, thereby offering complementary evidence to trial-based literature. Despite its strengths, this study has notable limitations. Retrospective design inherently risks selection bias, as patients undergoing regional anesthesia may represent less complex surgical cases or distinct comorbidity profiles compared with those receiving general anesthesia. Although covariate adjustments were applied, unmeasured confounding, such as intraoperative management practices or anesthesiologist preference, remains possible. Group size imbalance further constrained analysis, with approximately 24,545 patients in the general anesthesia group but only 127 in the regional anesthesia group; estimates for the regional arm are therefore less precise.

Moreover, reliance on the MIMIC-IV dataset, drawn from a single U.S. tertiary hospital, restricts the generalizability of these findings to broader healthcare contexts. The limitation that databases like MIMIC-IV do not capture postoperative functional outcomes such as long-term pain, cognitive recovery, or quality of life is

increasingly acknowledged in anesthesia research. Functional outcomes and QoR after anesthesia and surgery are complex, multidimensional processes that encompass physical, psychological, and social domains, which traditional datasets often fail to capture (27, 28). The QoR scales, such as the QoR-15 and QoR-40, provide patient-centered metrics but are not routinely included in large clinical databases, thereby limiting comprehensive outcome assessment (29).

Several mechanisms may explain why regional anesthesia was associated with improved outcomes. First, regional anesthesia may reduce sympathetic activation and blunt the surgical stress response, thereby contributing to more stable cardiovascular physiology (30). Second, by reducing the need for tracheal intubation and mechanical ventilation in selected procedures, regional anesthesia may lower the risk of pulmonary complications such as pneumonia and ventilator-associated lung injury (31). Third, regional approaches may attenuate systemic inflammation (e.g., lower perioperative cytokine release), potentially reducing the risk of infection and sepsis (32). In older adults, regional anesthesia has been associated with a lower incidence of postoperative delirium and cognitive dysfunction compared with general anesthesia in some studies, which may translate into more favorable recovery trajectories (33-35).

From a clinical perspective, these findings suggest that anesthetic technique should be considered not only from a procedural standpoint but also as part of individualized perioperative risk management. In frail or multimorbid patients, regional anesthesia may offer meaningful reductions in morbidity and resource utilization when feasible; conversely, general anesthesia remains essential for complex or long-duration operations where regional techniques are impractical or contraindicated. Practical considerations, including surgical site, expected duration, anticoagulation status, patient preference, and operator expertise, as well as the possibility of block failure and conversion to general anesthesia, should be incorporated into shared decision-making. Machine-learning tools, such as those evaluated here, could eventually support clinicians by identifying patients most likely to benefit from regional approaches, pending external validation, calibration, and decision-curve analysis. They should not complement clinical judgment.

Further research is needed through multicenter prospective studies to confirm these associations in broader populations. Randomized controlled trials remain crucial for establishing causality, particularly for multimodal anesthesia strategies and novel agents such as remimazolam and ciprofol. Long-term endpoints, including persistent pain, cognitive recovery, and quality of life, should be incorporated. Prospective registries should also include patient-reported outcomes to capture QoR more comprehensively. Ultimately, evaluating AI-enabled perioperative decision support in pragmatic trials could clarify its impact on patient safety, efficiency, and outcomes.

### 5.1. Conclusions

Our findings from a retrospective analysis of more than 30,000 patients in the MIMIC-IV database indicate that the choice of anesthesia has a measurable impact on perioperative outcomes, with regional anesthesia associated with lower in-hospital mortality, reduced infection rates, and shorter hospital and ICU stays compared with general anesthesia. By incorporating ML methods, particularly the XGBoost algorithm, we also demonstrated that anesthesia type can be predicted with high accuracy using routine demographic and clinical features, highlighting the potential of predictive analytics to support personalized anesthesia planning. While the retrospective design, imbalance in group sizes, and lack of long-term functional outcomes limit causal inference, the consistency of associations across multiple endpoints underscores the clinical importance of anesthesia modality as more than a technical consideration, but rather as a determinant of patient recovery and healthcare resource use. Future prospective multicenter studies and randomized controlled trials are needed to confirm these observations, integrate long-term and patient-reported outcomes, and evaluate emerging agents and multimodal strategies. The use of AI in perioperative decision-making may further enhance individualized and patient-centered care in anesthesia.

### 5.2. Limitations

This retrospective, observational study precludes causal inference; findings should be interpreted as associations. Although we adjusted for measured covariates, residual and unmeasured confounding (e.g.,

intraoperative management, provider preference) may remain. Procedure information was available only at a coarse level; detailed factors such as urgency, complexity, surgical approach, and ASA class were not comprehensively captured, which may influence both anesthetic selection and outcomes (confounding by indication). We used secondary data from MIMIC-IV (single U.S. tertiary center), limiting generalizability; registry data can include missingness, miscoding, and exposure misclassification (e.g., combined general-plus-regional techniques labeled as a single category). Preprocessing choices (e.g., imputation and outlier handling) may affect estimates and should be viewed as analytic assumptions. There was marked class imbalance (general  $\gg$  regional); we addressed this only in prediction models using SMOTE within training folds to avoid leakage, but the small regional cohort still reduces the precision/stability of estimates for that group. For ICU time-to-discharge, time-to-event analyses may be affected by competing risks and time-dependent confounding. Finally, while machine-learning models performed well internally, external validation, calibration, and decision-curve analysis are needed before clinical deployment.

## Footnotes

**Authors' Contribution:** Study concept and design: G. A., E. N., and T. A.; Acquisition of data: G. A.; Analysis and interpretation of data: G. A. and A. A.; Drafting of the manuscript: P. V., J. M. and Y. T.; Critical revision of the manuscript for important intellectual content: P. V. and Y. T.; Statistical analysis: G. A. and A. A.; Administrative, technical, and material support: G. A., T. A., and E. N.; Study supervision: E. N. and T. A.

**Conflict of Interests Statement:** The authors declare no conflict of interest.

**Data Availability:** The MIMIC-IV database is a publicly available, de-identified dataset.

**Funding/Support:** This study was funded by Shahid Beheshti University of Medical Sciences.

## References

- Gustafsson UO, Oppelstrup H, Thorell A, Nygren J, Ljungqvist O. Adherence to the ERAS protocol is Associated with 5-Year Survival After Colorectal Cancer Surgery: A Retrospective Cohort Study. *World J Surg.* 2016;**40**(7):1741-7. [PubMed ID: 26913728]. <https://doi.org/10.1007/s00268-016-3460-y>.
- Ripolles-Melchor J, Ramirez-Rodriguez JM, Casans-Frances R, Aldecoa C, Abad-Motos A, Logrono-Egea M, et al. Association Between Use of Enhanced Recovery After Surgery Protocol and Postoperative Complications in Colorectal Surgery: The Postoperative Outcomes Within Enhanced Recovery After Surgery Protocol (POWER) Study. *JAMA Surg.* 2019;**154**(8):725-36. [PubMed ID: 31066889]. [PubMed Central ID: PMC6506896]. <https://doi.org/10.1001/jamasurg.2019.0995>.
- Raaja GMT. Impact of Anesthesia on Postoperative Recovery: A Comparative Study. *Int J Pharm Clin Res.* 2023;**15**(6):2225-38.
- John Doyle D, Dahaba AA, LeManach Y. Advances in anesthesia technology are improving patient care, but many challenges remain. *BMC Anesthesiology.* 2018;**18**(1). <https://doi.org/10.1186/s12871-018-0504-x>.
- Seger C, Cannesson M. Recent advances in the technology of anesthesia. *Fl000Res.* 2020;**9**. [PubMed ID: 32494358]. [PubMed Central ID: PMC7236591]. <https://doi.org/10.12688/fl000research.24059.1>.
- Fu G, Xu L, Chen H, Lin J. State-of-the-art anesthesia practices: a comprehensive review on optimizing patient safety and recovery. *BMC Surg.* 2025;**25**(1):32. [PubMed ID: 39833810]. [PubMed Central ID: PMC11749226]. <https://doi.org/10.1186/s12893-025-02763-6>.
- Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med.* 2019;**380**(14):1347-58. [PubMed ID: 30943338]. <https://doi.org/10.1056/NEJMr1814259>.
- Lauritsen SM, Kristensen M, Olsen MV, Larsen MS, Lauritsen KM, Jorgensen MJ, et al. Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nat Commun.* 2020;**11**(1):3852. [PubMed ID: 32737308]. [PubMed Central ID: PMC7395744]. <https://doi.org/10.1038/s41467-020-17431-x>.
- Xue B, Li D, Lu C, King CR, Wildes T, Avidan MS, et al. Use of Machine Learning to Develop and Evaluate Models Using Preoperative and Intraoperative Data to Identify Risks of Postoperative Complications. *JAMA Netw Open.* 2021;**4**(3). e212240. [PubMed ID: 33783520]. [PubMed Central ID: PMC8010590]. <https://doi.org/10.1001/jamanetworkopen.2021.2240>.
- Hyland SL, Faltys M, Huser M, Lyu X, Gumbsch T, Esteban C, et al. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med.* 2020;**26**(3):364-73. [PubMed ID: 32152583]. <https://doi.org/10.1038/s41591-020-0789-4>.
- Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. *JAMA.* 2018;**319**(13):1317-8. [PubMed ID: 29532063]. <https://doi.org/10.1001/jama.2017.18391>.
- Moody GB, Mark RG. A database to support development and evaluation of intelligent intensive care monitoring. *Computers in Cardiology 1996. IEEE; 1996.* p. 657-60.
- Saeed M, Villarroel M, Reisner AT, Clifford G, Lehman LW, Moody G, et al. Multiparameter Intelligent Monitoring in Intensive Care II: a public-access intensive care unit database. *Crit Care Med.* 2011;**39**(5):952-60. [PubMed ID: 21283005]. [PubMed Central ID: PMC3124312]. <https://doi.org/10.1097/CCM.0b013e31820a92c6>.
- Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data.*



- 2016;**3**:160035. [PubMed ID: [27219127](#)]. [PubMed Central ID: [PMC4878278](#)]. <https://doi.org/10.1038/sdata.2016.35>.
15. Johnson AE, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data*. 2023;**10**(1). <https://doi.org/10.1038/s41597-022-01899-x>.
16. Barbosa EC, Espirito Santo PA, Baraldo S, Meine GC. Remimazolam versus propofol for sedation in gastrointestinal endoscopic procedures: a systematic review and meta-analysis. *Br J Anaesth*. 2024;**132**(6):1219-29. [PubMed ID: [38443286](#)]. <https://doi.org/10.1016/j.bja.2024.02.005>.
17. Luo H, Tang Z. Efficacy and safety of remimazolam vs. propofol for general anesthesia with tracheal intubation: Systematic review and meta-analysis. *Biomed Rep*. 2025;**22**(1):13. [PubMed ID: [39624780](#)]. [PubMed Central ID: [PMC11609608](#)]. <https://doi.org/10.3892/br.2024.1891>.
18. Roh YH, Park SG, Lee SH. Regional versus General Anesthesia in Postoperative Pain Management after Distal Radius Fracture Surgery: Meta-Analysis of Randomized Controlled Trials. *J Pers Med*. 2023;**13**(11). [PubMed ID: [38003859](#)]. [PubMed Central ID: [PMC10671853](#)]. <https://doi.org/10.3390/jpm13111543>.
19. Verret M, Le JBP, Lalu MM, Jeffers MS, McIsaac DI, Nicholls SG, et al. Effectiveness of dexmedetomidine on patient-centred outcomes in surgical patients: a systematic review and Bayesian meta-analysis. *Br J Anaesth*. 2024;**133**(3):615-27. [PubMed ID: [39019769](#)]. [PubMed Central ID: [PMC11347795](#)]. <https://doi.org/10.1016/j.bja.2024.06.007>.
20. Amoroso K, Okano I, Sarin M, Hughes AP, Zelenty WD, Shue J, et al. Comparative effectiveness of anesthetic technique on outcomes after lumbar spine surgery: a retrospective propensity score-matched analysis of the National Surgical Quality Improvement Program, 2009-2019. *Reg Anesth Pain Med*. 2023;**48**(7):343-8. [PubMed ID: [36650026](#)]. <https://doi.org/10.1136/rapm-2022-104098>.
21. Shi X, Zhang J, Sun Y, Chen M, Han F. Effect of different sedatives on the prognosis of patients with mechanical ventilation: a retrospective cohort study based on MIMIC-IV database. *Front Pharmacol*. 2024;**15**:1301451. [PubMed ID: [39092229](#)]. [PubMed Central ID: [PMC11291308](#)]. <https://doi.org/10.3389/fphar.2024.1301451>.
22. Park S, Choi AY, Park E, Park HJ, Lee J, Lee H, et al. Effects of continuous ketamine infusion on hemodynamics and mortality in critically ill children. *PLoS One*. 2019;**14**(10). e0224035. [PubMed ID: [31626685](#)]. [PubMed Central ID: [PMC6799949](#)]. <https://doi.org/10.1371/journal.pone.0224035>.
23. Hudaib M, Malik H, Zakir SJ, Rabbani S, Gnanendran D, Syed ARS, et al. Efficacy and safety of ciprofol versus propofol for induction and maintenance of general anesthesia: a systematic review and meta-analysis. *J Anesth Analg Crit Care*. 2024;**4**(1):25. [PubMed ID: [38605424](#)]. [PubMed Central ID: [PMC11008023](#)]. <https://doi.org/10.1186/s44158-024-00160-8>.
24. Handley GH, Silbert BS, Mooney PH, Schweitzer SA, Allen NB. Combined general and epidural anesthesia versus general anesthesia for major abdominal surgery: postanesthesia recovery characteristics. *Reg Anesth*. 1997;**22**(5):435-41. [PubMed ID: [9338905](#)]. [https://doi.org/10.1016/s1098-7339\(97\)80030-2](https://doi.org/10.1016/s1098-7339(97)80030-2).
25. Wang J, Zhang W, Qian J, Zhong Y, Deng Y, Leng W, et al. Effects of combined anesthesia on pulmonary oxygenation function, hemodynamics, and respiratory compliance in elderly patients undergoing pulmonary lobectomy for lung cancer. *Medicine (Baltimore)*. 2024;**103**(45). e40325. [PubMed ID: [39533631](#)]. [PubMed Central ID: [PMC11557111](#)]. <https://doi.org/10.1097/MD.00000000000040325>.
26. Hung K, Chen J, Wu S, Huang P, Wu J, Liu T, et al. meta-analysis comparing the efficacy and safety of ciprofol (HSK3486) versus propofol for. *Advances in Novel Pharmacotherapeutics and Drug Discovery: Computational, Experimental, Translational, and Clinical Models*. 2024.
27. Sreepada RS, Chang AC, West NC, Sujan J, Lai B, Poznikoff AK, et al. Dashboard of Short-Term Postoperative Patient Outcomes for Anesthesiologists: Development and Preliminary Evaluation. *JMIR Perioper Med*. 2023;**6**. e47398. [PubMed ID: [37725426](#)]. [PubMed Central ID: [PMC10548316](#)]. <https://doi.org/10.2196/47398>.
28. Leger M, Campfort M, Cayla C, Lasocki S, Rineau E. Postoperative quality of recovery measurements as endpoints in comparative anaesthesia studies: a systematic review. *Br J Anaesth*. 2021;**126**(6):e210-2. [PubMed ID: [33892949](#)]. <https://doi.org/10.1016/j.bja.2021.03.008>.
29. Sharma R, Moied S, Raikwar S, Gupta V. Functional outcomes and quality of recovery after anaesthesia and surgery - Outreaching towards protracted goals. *Indian J Anaesth*. 2022;**66**(Suppl 3):S133-6. [PubMed ID: [35774242](#)]. [PubMed Central ID: [PMC9238233](#)]. [https://doi.org/10.4103/ija.ija\\_356\\_22](https://doi.org/10.4103/ija.ija_356_22).
30. Hutton M, Brull R, Macfarlane AJR. Regional anaesthesia and outcomes. *BJA Educ*. 2018;**18**(2):52-6. [PubMed ID: [33456810](#)]. [PubMed Central ID: [PMC7807931](#)]. <https://doi.org/10.1016/j.bjae.2017.10.002>.
31. Hausman MJ, Jewell ES, Engoren M. Regional versus general anesthesia in surgical patients with chronic obstructive pulmonary disease: does avoiding general anesthesia reduce the risk of postoperative complications? *Anesth Analg*. 2015;**120**(6):1405-12. [PubMed ID: [25526396](#)]. <https://doi.org/10.1213/ANE.0000000000000574>.
32. Cruz FF, Rocco PR, Pelosi P. Anti-inflammatory properties of anesthetic agents. *Crit Care*. 2017;**21**(1):67. [PubMed ID: [28320449](#)]. [PubMed Central ID: [PMC5359894](#)]. <https://doi.org/10.1186/s13054-017-1645-x>.
33. Yang X, Wang Y, Qian F, Lu L, Bi D, Tung T. Risk of dementia in older patients with different anesthesia: a systematic review and meta-analysis of cohort studies. *BMC Geriatrics*. 2025;**25**(1). <https://doi.org/10.1186/s12877-025-06096-0>.
34. Chen IW, Sun CK, Chen JY, Chen HT, Lan KM, Hung KC, et al. Comparison of regional vs. general anesthesia on the risk of dementia: a systematic review and meta-analysis. *Front Public Health*. 2024;**12**:1362461. [PubMed ID: [38887243](#)]. [PubMed Central ID: [PMC1182446](#)]. <https://doi.org/10.3389/fpubh.2024.1362461>.
35. Sprung J, Schulte PJ, Knopman DS, Mielke MM, Petersen RC, Weingarten TN, et al. Cognitive function after surgery with regional or general anesthesia: A population-based study. *Alzheimers Dement*. 2019;**15**(10):1243-52. [PubMed ID: [31495602](#)]. [PubMed Central ID: [PMC6943821](#)]. <https://doi.org/10.1016/j.jalz.2019.06.4949>.