J Cell Mol Anesth. 2024 March; 9(1): e145369. https://doi.org/10.5812/jcma-145369.

Published online: 2024 February 3. **Review Article**

Machine Learning-Guided Anesthesiology: A Review of Recent Advances and Clinical Applications

Sana Hashemi [1], Zohreh Yousefzadeh [1], Ahmad Ali Abin [1,*], Azar Ejmalian [2], Shahabedin Nabavi [1], Ali Dabbagh [3]

[1] Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran
[2] Department of Anesthesiology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[3] Anesthesiology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[*] *Corresponding author:* Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran. PO Box: 1983969411, Tel: +989127253488, Email: a_abin@sbu.ac.ir

Received 2024 January 22; **Accepted** 2024 February 3.

Abstract

Anesthesia is the process of inducing and experiencing various conditions, such as painlessness, immobility, and amnesia, to facilitate surgeries and other medical procedures. During the administration of anesthesia, anesthesiologists face critical decision-making moments, considering the significance of the procedure and potential complications resulting from anesthesia-related choices. In recent years, artificial intelligence (AI) has emerged as a supportive tool for anesthesia decisions, given its potential to assist with control and management tasks. This study aims to conduct a comprehensive review of articles on the intersection of AI and anesthesia. A review was conducted by searching PubMed for peer-reviewed articles published between 2020 and early 2022, using keywords related to anesthesia and AI. The articles were categorized into nine distinct groups: “Depth of anesthesia", “Control of anesthesia delivery", “Control of mechanical ventilation and weaning", “Event prediction", “Ultrasound guidance", “Pain management", “Operating room logistic", “Monitoring", and “Neuro-critical care". Four reviewers meticulously examined the selected articles to extract relevant information. The studies within each category were reviewed by considering items such as the purpose and type of anesthesia, AI algorithms, dataset, data accessibility, and evaluation criteria. To enhance clarity, each category was analyzed with a higher resolution than previous review articles, providing readers with key points, limitations, and potential areas for future research to facilitate a better understanding of each concept. The advancements in AI techniques hold promise in significantly enhancing anesthesia practices and improving the overall experience for anesthesiologists.

Keywords: Anesthesia, Artificial Intelligence, Machine Learning, Decision-Making, Predictive Model

1. Introduction

Currently, artificial intelligence (AI) is becoming increasingly pervasive across various scientific fields, including medicine and healthcare. Medical applications of AI have seen significant growth in recent years (1, 2). Given the direct impact of medical decisions and activities on human life, the healthcare sector receives substantial attention from the research community. In this context, AI's potential to assist physicians and medical staff in managing complex tasks, handling a body of data, and making medical decisions has garnered considerable interest.

Numerous studies conducted in recent years have demonstrated the high capabilities of AI algorithms, resulting in a noticeable reduction of risks associated with medical practices. Furthermore, AI implementation has contributed to providing a better overall experience for both patients and medical staff alike.

The field of anesthesia is one of the critical areas in medical sciences. Since conscious patients are unable to breathe due to anesthetic drugs, the anesthesiologist must ensure they have stable breathing conditions through appropriate interventions. Therefore, making accurate decisions in this context is of utmost importance. Tasks such as predicting the depth of intraoperative anesthesia (3), developing metrics for neurological care (4), and predicting postoperative complications (5) are among the responsibilities that anesthesiologists perform with the assistance of AI algorithms. This integration of AI results in reduced errors, increased speed, and enhanced accuracy, as reported by anesthesiologists.

This study is a review of current trends in AI in anesthesia within 2020-2022. Moreover, this study provides substantial assistance for anesthesiologists who are familiar with the basic concepts of AI and machine learning (ML) and are interested in monitoring the latest developments in this field. Instead of focusing extensively on teaching the basics of ML and AI, the authors have dedicated their efforts to reviewing existing studies and covering a wider range of articles.

After reviewing the studies in the field of anesthesia and AI, existing studies were summarized by presenting a tabular structure of the main details within each study. This structure includes information on the purpose and type of surgery and anesthesia, the AI algorithms employed, the features used by AI models, dataset information, data accessibility, and evaluation criteria. Additionally, this study categorized and reviewed articles with a higher level of detail than Hashimoto et al.'s study (6), which divided the studies into six categories. This study extends the previous research by introducing three new categories and incorporating recent studies from 2020. The collected studies were thoughtfully divided into nine categories based on the expertise of the anesthesiologist (Table 1).

The remainder of the study is structured as follows: Section 2 provides an overview of AI and ML techniques. The categorization and reviews are presented in Section 3. A comprehensive discussion of the reviewed studies is presented in Section 4. Finally, Section 5 concludes the study and offers insights into potential future directions.

2. A Brief Introduction to AI and ML

Artificial intelligence is a branch of computer science whose main purpose is to produce intelligent machines capable of performing tasks that require human intelligence. This technology is a type of human intelligence simulation for computers, mainly aiming to design and build machines that can think like humans and imitate their behavior. Artificial intelligence techniques can be divided into several major categories, and currently, the two categories of ML and deep learning (DL) are widely used in various applications. This section provides a brief description of the concepts of ML and DL and the interpretability of learning models for the reader's general acquaintance.

Machine learning is considered one of the most important branches of AI. In ML, the learning process begins with observations in the form of data. The learner uses examples, direct experiences, or instructions to identify specific patterns and automatically make decisions and solve problems. Machine learning algorithms are typically categorized based on their learning styles, such as supervised learning, unsupervised learning (7), and semi-supervised learning, depending on the observability of variables under investigation.

Deep learning is a subset of ML that mimics the way the human mind learns about specific subjects. Deep learning aims to learn complex patterns by finding representations that fit each problem through successive layers of neural networks. Feature extraction is a key aspect of both ML and DL; however, DL algorithms are more automated than ML, where human resources might be involved in feature selection.

The interpretability of learning model outcomes is one of the most important issues in both ML and DL. When considering a particular medical problem, a learning algorithm can inform the physician of various predictions related to the problem. However, it might not provide the physician with sufficient information about the underlying reasons for those predictions and the process of reaching them. This "black box" nature of the learning algorithm might limit its applications in the medical field. To address this challenge, interpretability techniques are applied to the immediate results of the models. A model is considered interpretable when one can easily and significantly grasp the reasoning behind its predictions and decisions. More interpretable models are easier for human resources to understand and trust, especially in critical domains, such as healthcare (8).

3. Literature Review

This section studies all articles within a specific category based on their similarities.

3.1. Category A: Neuro-critical Care

The brain is the most vital organ of the human body. Therefore, specialists must pay special attention to brain function during anesthesia. Managing and controlling the function of the brain and other organs is done through neuro-critical care. Predicting and monitoring brain damage can be challenging for human resources, and as a result, AI algorithms have been used to create systems for performing such tasks. Brain injuries can be divided into two groups: Traumatic and non-traumatic.

Table 1. Article Categories with the Number of Articles in Each Category

Category	Names		#Articles
A	Neuro-critical care		9
B	Pain management		6
C	Control of mechanical ventilation and weaning		2
D	Event prediction	D1. Perioperative	15
		D2. Postoperative	16
		D3. Critical care	12
E	Ultrasound guidance		5
F	Operating room logistic		3
G	Depth of anesthesia		21
H	Control of anesthesia delivery		6
I	Monitoring		4

Traumatic injuries result from trauma or brain injury; nevertheless, non-traumatic brain injuries are caused by vascular accidents, such as rupture or bleeding in the brain or narrowing of the arteries (cerebral ischemia). Most studies in this category are related to traumatic brain and head injuries.

Intracranial hemorrhage is considered one of the most traumatic brain injuries. In two studies (9) and (10), deep neural networks and unsupervised ML algorithms were employed to analyze this injury, respectively. In another study by Schweingruber et al. (4), a deep, long, short-term memory (LSTM) neural network was used to predict the critical stages of intracranial hypotension and intracranial pressure, which are types of traumatic brain injuries. Farzaneh et al. (11, 12) conducted two studies focused on the use of AI methods to classify and predict different types of brain damage. In 2020, they used an ML model to assess the severity of subdural hematoma (11). In 2021, they provided long-term performance outcomes for patients with traumatic brain injury (TBI) by presenting an ML framework (12). The latter study's results were interpreted using the Shapley method.

Seizures are bursts of uncontrolled electrical activity between brain cells, causing temporary abnormalities in the function of some organs. The random forest (RF) model was used to diagnose and monitor traumatic brain injuries related to seizures in a study (13); however, another study (14) used the generalized linear model (GLM). Both studies utilized continuous electroencephalogram (EEG) signals. The use of interpretability techniques is one of the advantages and positive contributions of these studies. Given the importance of interpretability in medical applications, as mentioned in section 3, it is noteworthy that its significance in the field of neurological care is amplified due to the presence of the most critical organ in the human body. Among the traumatic injuries of the head, subarachnoid hemorrhage (SAH) injuries were predicted by Koch et al. (15) using the Elastic-Net ML model and orthogonal partial least squares-discriminant analysis (OPLS-DA). Hypoxic-ischemic brain injury is a non-traumatic brain injury that can occur after cardiac arrest. In another study, Elmer et al. developed a new clustering algorithm called K-prototypes, inspired by the famous K-means clustering algorithm, to identify the phenotypes of primary brain damage after cardiac arrest (16). Further details about the reviewed studies in this category are shown in Table 2.

3.2. Category B: Pain Management

Pain is a sensation caused by stimulating nociceptors in the central or peripheral nervous system. This feeling can arise following a surgical incision and might result from inadequate anesthetic drug injection during an operation or insufficient postoperative analgesia. Therefore, pain management and prevention are of great importance for both the anesthesiologist and the patient. Artificial intelligence models can assist specialists in better pain management by measures, such as defining the pain index and predicting its timing. The reviewed studies in this category can be divided into two subcategories.

The first subcategory includes studies aiming to diagnose and predict the occurrence of pain during or after an operation. For example, in a study (13), the possibility of diagnosing toothache based on the three signals of electrocardiography (ECG), photoplethysmography (PPG), and chest were investigated using the RF model, which performed well on the test dataset. Tan et al. (17) compared ML techniques to statistical inference techniques to identify and predict breakthrough pain during labor, with the

ML models not performing better than the statistical methods, potentially due to the presence of unbalanced data.

The second subcategory of studies focuses on assessing the level of pain by defining a pain index. The lack of a well-defined criterion for determining and measuring a patient's pain level to adapt drug injections during general anesthesia is a major challenge. Gonzalez-Cava et al. (18) aimed to evaluate the performance of the pain index using ML classifiers; however, another study (19) indicated that monitoring the injectable drug dose using the pain level index helped reduce postoperative pain. The Nociception level (NOL) index is another multi-parameter AI-based index designed to monitor pain during general anesthesia, which was observed to reduce postoperative pain. In another study (20), a new relief index was developed using photoplethysmogram spectroscopy and a convolutional neural network (CNN) to assess pain in conscious patients.

Rebound pain is a common outcome that occurs after a peripheral nerve block, usually subsiding 24 to 48 hours after the block was formed, often occurring after outpatient operations for patients. To address this issue, Barry et al. (21) used ML models to examine factors associated with rebound pain in patients who received peripheral nerve blocks for outpatient operations. A metric called the numerical rating scale (NRS) was defined to measure the level of pain in this study. The logistic model tree attribute-selected classifier with receiver operating characteristic (ROC) showed the best-reported result at around 60%. Table 3 shows the main points of the reviewed studies in this category.

3.3. Category C: Control of Mechanical Ventilation and Weaning

Mechanical ventilation is a life-supporting treatment that aids patients who are unable to breathe on their own. It involves the use of a mechanical device, such as a ventilator, artificial respiration device, or respiratory system, to assist patients in breathing. Patients requiring respiratory support due to a serious illness are typically hospitalized in the intensive care unit (ICU). However, mechanical ventilation can pose challenges, such as patient restlessness caused by the use of lighter anesthesia and inadequate oxygen supply to the respiratory organs. Artificial intelligence models have been employed to address these challenges effectively.

Two reviewed studies in this category utilized ML algorithms to predict and manage challenges related to patient restlessness due to lighter anesthesia and insufficient oxygen supply to the respiratory organs. The use of lighter sedatives with lighter anesthetics is often recommended to improve aggressive mechanical ventilation, reduce mortality, and enhance clinical outcomes. However, this approach can lead to issues, such as accidental extubation and patient-ventilator asynchrony. Additionally, the use of lighter sedatives might increase the risk of patient agitation in response to other nervous stimulation. To tackle these challenges, timely prediction of patient agitations and their management is crucial when using lighter anesthesia. Therefore, one study (23) developed a collective ML model to predict patient agitation in the ICU over the next 24 hours.

Another significant aspect of mechanical ventilation is assessing spontaneous breathing (SB) attempts, which is an essential criterion in respiratory drive. However, SB levels can vary due to various factors, including evolving pathology and sedation levels. Therefore, the continuous assessment of SB is necessary. In a study (24), a convolutional autoencoder (CAE) was developed to quantify the amount of SB using airway pressure and flow waveform data. The characteristics of each reviewed study in this category are summarized in Table 4.

3.4. Category D: Event Prediction

This category examines studies that aim to predict events, which involve estimating the probability of specific occurrences in the future. Artificial intelligence algorithms have been employed as tools to enhance the accuracy, ease, and speed of predicting these events and preventing related complications. The events are categorized into three subcategories: Perioperative, postoperative, and critical care, each of which will be discussed in more detail below.

3.4.1. Subcategory D1: Perioperative

Perioperative events refer to occurrences that might happen to a patient before, during, or immediately after an operation. In this subcategory, the prediction of such events is the focus (25, 26). A common perioperative event is fluctuations in blood pressure, particularly hypotension, which can lead to serious complications, such as cardiovascular injury or even death. Several articles in this subcategory predicted hypotension before its occurrence to enable specialists to take necessary tasks to prevent it (27-33). Another crucial event is difficult laryngoscopy, defined as the inability to visualize part of the vocal cords during multiple laryngoscopy attempts by a trained anesthesiologist. Predictive models for difficult laryngoscopy were

Table 4. Category C: Control of Mechanical Ventilation and Weaning

No.	Study	Goal	Type of Anesthesia	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance	Interpretable?
1	Zhang et al. (2021) (23)	Prediction of agitation in invasive mechanical ventilation patients under light sedation.	Sedation	Unavailable	578/Some ICUs in 80 Chinese hospitals	Risk factors for delirium identified, ventilator parameters that can influence asynchrony, including ventilation mode, positive end-expiratory pressure, plateau pressure, Fio2, respiratory rate, and minute ventilation	Adaboost, Linear SVM with Class Weights, C5.0, XGboost, An ensemble model including four mentioned models	Ensemble model	AUC: 0.918	Yes, using the "BreakDown" algorithm
2	Ang et al. (2021) (24)	To quantify the magnitude of spontaneous breathing (SB) effort using only bedside (mechanical ventilation) MV airway pressure and flow waveform	-	Unavailable	13.6M+1800/simulated SB flow and normal flow data (NB)+National University of Singapore Hospital (test data)	SB flow	Convolutional autoencoder	Convolutional auto encoder	MSE: 4.77	No

Abbreviations: ICU, intensive care unit; SVM, support vector machine; XGboost, extreme gradient boosting; AUC, area under the curve; SB, spontaneous breathing; NB, normal breathing; MSE, mean square error.

developed ML techniques in the studies of this subcategory (34, 35).

Additionally, Mathis et al. (36) utilized ML approaches to identify patients who ultimately faced postoperative heart failure with reduced ejection fraction (HFrEF). The aforementioned study demonstrated that the extreme gradient boosting algorithm outperformed other ML algorithms in this prediction task. Other studies in this subcategory applied DL to improve the detection of life-threatening arrhythmia (37), classify ECG signals for anesthesia assessment (38), and investigate the elements of synaptic transmission based on the anesthetized patient's EEG data (39).

3.4.2. Subcategory D2: Postoperative

The postoperative period encompasses events occurring at long intervals after an operation (40, 41). In most of the reviewed articles in this subcategory, predicted events include postoperative complications in specific conditions or diseases.

Postoperative delirium was predicted in three studies using ML algorithms (5, 42). In addition to predicting delirium, several studies in this subcategory utilized ML algorithms to predict blood pressure fluctuations during the postoperative period. Palla et al. (43) and Schenk et al. (44) predicted postoperative hypotension; however, another study predicted an increase in postoperative hypertension (45). Other studies used ML techniques to predict postoperative complications, such as cardiac events (46), cerebral infarction and myocardial infarction (47), and acute kidney injury (48). Cao et al. (49) employed DL algorithms to predict serious complications after bariatric surgery. Qian et al. (50) presented a study evaluating the importance of operation time in classifying surgical complications using interpretable ML approaches.

Moreover, one study (41) introduced a tool called the surgical and medical postoperative complications prediction tool (SUMPOT) based on an artificial neural network to identify patients at risk of postoperative complications. Additionally, the relationship between cannabis use and a slight increase in the risk of postoperative nausea and vomiting was investigated using ML (51). Two studies in 2021 by Lu et al. (52, 53) focused on identifying patients in need of anterior cruciate ligament reconstruction (ACLR) (52) and predicting the cost of ACLR (53).

3.4.3. Subcategory D3: Critical Care

Studies in this subcategory primarily focused on predicting events related to clinical interventions for patients frequently admitted to the ICU. The worst postoperative event in this subcategory is patient death (54, 55). Several studies, similar to the previous subcategories, predicted hypotension by considering clinical interventions. Cherifa et al. (56) predicted hypotension in the ICU using deep neural networks; nevertheless, two other studies employed different ML algorithms for the same prediction (57, 58). Additionally, Hu et al. (59) used ML techniques to develop a model for predicting seizures in critically ill children. Myasthenia gravis (MG), a neuromuscular disorder associated with acquired autoimmunity causing muscle weakness, was also investigated in this subcategory, where Chang et al. (60) developed a decision tree-based model to predict the severity of MG.

Furthermore, predicting tracheal intubations was considered crucial in the ICU, especially for medical personnel not familiar with the procedure. Hayasaka et al. (61) designed an AI model using a CNN to classify difficult intubations based on the patient's facial image. Machine learning methods and statistical techniques were also used to investigate the relationship between positive cultures during hospitalization and long-term outcomes in critically ill surgical patients (62), the relationship between red cell distribution width (RDW) and prognosis in patients with sepsis-associated thrombocytopenia (SAT) (63), and the relationship between primary brain magnetic resonance imaging (MRI) data and functional outcomes of patients with severe herpes simplex encephalitis (HSE) 90 days after ICU admission (64). Moreover, a study (65) explored parametric and non-parametric methods for predicting cerebral performance category (CPC) using longitudinal data after cardiac arrest. Further detailed information about the reviewed studies in this category can be found in Tables 5, 6, and 7.

3.5. Category E: Ultrasound Guidance

Determining the appropriate site for injecting an anesthetic drug is a significant challenge in anesthesia, particularly in regional anesthesia. Injecting the drug around the relevant nerve is essential to achieve nerve block, temporarily blocking pain signals. However, injecting the drug at the wrong site or at a long distance from the nerve can lead to dangerous complications. Anesthesiologists often face difficulty in accurately performing this task in real-time. To address this challenge, AI techniques, particularly image processing, have been employed to induce regional anesthesia under ultrasound guidance. These techniques allow physicians to visualize the internal structure of organs and determine the correct injection site more easily.

The reviewed studies in this category can be divided into two subcategories based on the AI algorithms used for determining the appropriate injection site: DL-based algorithms and tracking algorithms based on correlation filters.

Most of the articles in this category fall into the first subcategory. In one study (66), a novel algorithm was proposed for accurate needle tip placement under ultrasound guidance when the needle body is invisible and the tip has low intensity. The algorithm first extracts the needle tip properties in successive ultrasound frames using a detection scheme and then predicts the location of the needle tip using a deep neural network consisting of CNN and LSTM recurrent units. The study achieves an error rate of 0.06 ± 0.02 mm for the needle entry point and a processing time of 0.064 seconds. However, the limitations included using ex vivo data and specific needle types.

In another study (67), the DL model was used to determine the anesthesia site by dividing the patients into control and algorithm groups. The algorithm group used ultrasound guidance and a deep CNN SegNet (68) to determine the anesthesia site, leading to significant improvements in the average injection duration and needle insertion depth, compared to the control group.

In another study (69), a preliminary assessment of an AI system was performed using a deep CNN network for semantic segmentation of ultrasound images. The aforementioned study focused on seven specific nerve blocks, and the proposed model aimed to detect the presence of these seven nerve blocks in the input images.

Studies in the second subgroup focus on tracking arteries instead of nerves in ultrasound images due to the low quality of the images, making nerve detection difficult. In one study (70), real-time tracking models were designed using a modified kernelized correlation filter (KCF) and modified discriminative correlation filter with channel and spatial reliability method (CSR-DST). The CSR-DST algorithm performed faster; however, the KFC provided better results and was identified as the superior algorithm. Table 8 shows the key characteristics of the reviewed studies in this category.

3.6. Category F: Operating Room Logistics

The studies conducted in this category focused on organizing and coordinating the affairs within the operating room. Some of the studies in this category

Table 8. Category E: Ultrasound Guidance

No.	Study	Goal	Type of Anesthesia	Enhancement Filter(s)	Nerve Block(s)	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance
1	Paris and Hafiane (2021) (70)	To track arteries in ultrasound guidance to find a proper place to inject the anesthetic drugs	Regional	Kernelized Correlation filter, Discriminative Correlation filter	-	Unavailable	71/not reported	Fd is introduced as features that are extracted from images after applying the kernels.	Gradient descent applied to the search for ellipses, Modified KCF, Modified CSR-DST.	Modified CSR-DST	Mean Error: 15.16 STD Error: 25.51 FPS: 63.06 Precision ~ 95%
2	Bowness et al. (2021) (69)	To perform semantic segmentation of the input ultrasound videos	Regional	-	Supraclavicular level brachial plexus: Subclavian artery, brachial plexus nerves, first rib, pleura. Erector spinae plane (thoracic region): Trapezius/rhomboid/erector spinae (group) muscles, vertebral transverse process/rib, pleura. Rectus sheath: Rectus abdominis muscle, rectus sheath, peritoneal contents. Adductor canal: Femoral artery, saphenous nerve, sartorius/adductor longus, femur	Unavailable	144 & 244/The Royal Gwent Hospital, Ystrad Mynach Hospital, StWoolos Hospital & Nevill Hall Hospital	Extracted features from Deep CNN	Deep CNN Based on U-Net	Deep CNN Based on U-Net	Using statistical analysis, the Kruskal-Wallis H-test
3	Liu and Cheng (2021) (67)	To locate the anesthesia point of patients during scapular fracture surgery treated with the regional nerve block	Regional	Gaussian low-frequency filters	Scapula Regional Nerve Block	Request	100/Jiangxi Armed Police Corps Hospital	Ultrasound Images of the Scapula of the Patients	SegNet (A brand-new deep fully CNN)	SegNet	Injection Time: 7.7 ± 2.1 min Distance between the Puncture Point and the Scapula: 62.5 ± 7.2 mm
4	Mwikirize et al. (2021) (66)	Needle tip localization during challenging ultrasound-guided insertions when the shaft may be invisible, and the tip has a low-intensity	Regional	-	-	Unavailable	80/SonixGPS & Clarius C3	Enhanced Tip Images and B-Mode Images	DNN(CNN+LSTM)	DNN(CNN+LSTM)	Tip Localization Error: 0.52 ± 0.06 mm Overall Computation Time: 0.064 s

Abbreviations: KCF, kernelized correlation filter; CSR-DST, discriminative correlation filter with channel and spatial reliability method; FPS, frame per second; CNN, convolutional neural network; DNN, Deep Neural Network; LSTM, long short-term memory.

aimed to predict the duration of each operation (71, 72); however, others focused on addressing challenges that lead to the wastage of hospital facilities and resources (73). One significant challenge is day-of-surgery cancellation (DoSC), which can be problematic for hospital staff, patients, and their families, in addition to being costly and time-consuming. To address this issue, a study (73) analyzed the electronic file information of approximately 88 000 patients, considering various variables, including economic and social factors. The study utilized several ML algorithms to understand the reasons behind the DoSC.

In two other studies conducted by Gabriel et al. (71) and Jiao et al. (72), ML algorithms were used to predict the end time of surgery. Additionally, in Gabriel's study (71), predicting the patient's recovery period was another goal. All studies in this category utilized AI algorithms, particularly ML, to optimize hospital facilities and staff management. Table 9 shows further detailed information about the reviewed studies in this category.

3.7. Category G: Depth of Anesthesia

The anesthesia process consists of three stages: Anesthesia induction, maintenance of anesthesia, and recovery. In the anesthesia induction phase, the patient enters the initial phase of anesthesia when a specialist physician injects induction drugs, either through injection or inhalation. During the maintenance phase of anesthesia, the patient is maintained at an appropriate depth of anesthesia by administering the proper dose of maintenance medication. In the last stage, the patient recovers from anesthesia as the drugs are metabolized and eliminated from the body. Throughout these stages, the injection of relevant drugs by the anesthesiologist requires accurate knowledge and information about the depth of anesthesia and the patient's level of consciousness. Measuring the patient's physiological and clinical criteria simultaneously to assess the depth of anesthesia is challenging for physicians and prone to human errors. Artificial intelligence techniques can be employed to reduce these errors and improve performance in categorizing and monitoring the depth of anesthesia.

The studies in this category are divided into three groups based on the type of data used in each study. These groups include studies based on EEG signals,

Table 9. Category F: Operating Room Logistics

No.	Study	Goal	Type of Surgery	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance	Interpretable?
1	Gabriel et al. (2022) (71)	To predict the following composite outcome: 1. surgery finished by the end of the operating room block time and 2. the patient was discharged by the end of the recovery room nursing shift.	Outpatient surgery	Unavailable	13447/not reported	The surgical procedure, surgeon identification, ASA score, age, Gender, weight, surgical service line, scheduled surgical incision time, scheduled room time, actual room time, actual PACU length of stay	LR, RF Classifier, Balanced RF, Balanced Bagging, Simplefeedforward NN & SVM classifier	Balanced Bagging (Using SMOTE)	Precision: 83%; Recall: 77%; Matthew's correlation coefficient: 0.642; Sensitivity: 77.3%; Specificity: 87.1%; AUC: 0.905	Yes, the feature importance graph based on the balanced bagging approach
2	Jiao et al. (2020) (72)	To predict a continuous probability distribution of surgical case durations	Various surgical services	Unavailable	52735/Central operating location at St. Louis Children's Hospital, a free-standing, tertiary-care, pediatric hospital	Categorical (ASA, inpatient status, day of week), Continuous (scheduled surgery duration, patient age), Unstructured text (procedure name, surgical diagnosis) variables	A Neural Network (Mixture Density Network (MDN)), Tree-based methods (DT, RF, and GBT), non-ML statistical method (Bayesian statistical method)	MDN	Continuous Ranked Probability Score: 18.1 minutes	Yes, permutation importance was calculated for the MDN
3	Liu et al. (2021) (73)	To understand potential underlying contributors to disparities in DoSC rates across neighborhoods	-	Unavailable	88013/Cincinnati Children's Hospital Medical Center and Texas Children's Hospital	All features were in one of these categories: Transportation, Preoperative phone calls, Recent healthcare use, Prior cancellation behaviors, Surgery-related factors	Non-spatial regression models (GLM, L2-normalized GLM, SVM with polynomial kernels and DT, Spatial regression models (SAR model, spatial Durbin model, SEM, spatial Durbin error model, spatial moving average, and SAR confused models), CNNs & Graph Convolutional Networks	An L2-normalized generalized LR model	RMSE: 0.01305, 95% CI: 0.01257-0.01352	Yes, using feature importance generated from the best-performing L2-normalized generalized LR model

Abbreviations: ASA, American Society of Anesthesiology; PACU, Post Anesthesia Care Unit; LR, logistic regression; RF, random forest; NN, neural network; SVM, support vector machine; SMOTE, synthetic minority oversampling technique; AUC, area under the curve; MDN, mixture density network; DT, decision tree; GBT, gradient boosting-based tree; ML, machine learning; DoSC, day-of-surgery cancellation; GLM, generalized linear model; SAR, spatial autoregressive; SEM, spatial error model; CNN, convolutional neural network; RMSE, root mean square error.

physiological-clinical variables, and the combination of EEG signals and physiological-clinical variables.

The brain is the main human organ and the first area to be affected after injecting anesthetic drugs. Due to the good reflection of brain activity in EEG signals, they are used as supplement monitoring to determine the level of consciousness more accurately.

In studies based on EEG signals, researchers have developed monitoring systems using EEG-based criteria to evaluate the depth of anesthesia more accurately (3, 74-88). The bispectral index (BIS) is a common diagnostic index used to measure the depth of anesthesia based on EEG signals. In one study (75), a combined DL structure was proposed, consisting of three networks: CNNs using an inception module, LSTM, and one attention layer. The regression model's output was a BIS index used to determine the patient's depth of anesthesia, achieving 88.71% accuracy.

In other studies, new indices or improved versions of the previous indices were defined to determine the depth of anesthesia and enhance monitoring (78, 84). For instance, the Poincaré index was introduced to target a specific frequency range of 20 to 30 Hz, and it was combined with the classical Poincaré 0.5 - 47 Hz index using DL-improved anesthesia depth monitoring (84).

Using EEG-based indices as complementary monitoring can offer various benefits in assessing the depth of anesthesia and the patient's level of consciousness. However, there are certain limitations associated with EEG, such as low performance with volatile anesthetics, long latency, and susceptibility to interference from surgical stimulation. Apart from EEG-based studies, other data types have been used to train models and determine the level of consciousness in articles in this category (89-91). Dubost et al. (89) and Zhan et al. (90) utilized physiological or clinical and functional magnetic resonance imaging (fMRI) data as alternatives to EEG signals. Various methods, such as hidden Markov models and deep neural models, were employed as learning models in these studies (89, 90).

In the third group of studies, researchers combined EEG signals with other signals, such as auditory evoked potentials (AEP), to determine the level of consciousness by creating a new index (92, 93). In another study (93), several classification algorithms in ML were utilized to construct the unified index, with each model trained using EEG signal parameters as features. The support vector machine (SVM) model exhibited the best performance in this study, achieving a prediction probability of 0.935. The reviewed studies in this category are summarized in Table 10.

3.8. Category H: Control of Anesthesia Delivery

Managing the level of a patient's anesthesia by considering the appropriate dose of an anesthetic drug is a critical goal in the control of anesthesia delivery. Studies in this field focus on designing models to keep patients at certain levels of anesthesia or suggest measures to address challenges faced by anesthesiologists in this area. Making critical decisions about the patient's condition, such as determining the proper dose of an injectable anesthetic drug and controlling the patient's consciousness level by considering several parameters, is considered one of the most critical challenges for anesthesiologists during surgery. Some studies indicated that AI models could play an influential role in providing decision support for anesthesia delivery. Artificial intelligence algorithms can be utilized to determine the appropriate dose of an anesthetic drug for each patient, aiming to achieve the desired level of anesthesia.

In this category, many of the reviewed studies aim to determine or predict the appropriate dose of anesthetic drugs to achieve the desired level of anesthesia. For instance, Ingrande et al. (94) compared two biological models and a gated recurrent unit (GRU) network, where the GRU model demonstrated superior performance. Another study (95) designed a model to predict whether the patient will need remifentanil in the next n minutes using ML algorithms, such as SVM and the LSTM network, with LSTM being identified as the preferred model. Systolic blood pressure (SBP) was identified as the most important feature using the Shapley interpretability technique. In another study, Wei et al. (96) developed a decision tree model to determine the appropriate dose of local anesthetic hyperbaric bupivacaine during a cesarean section. The interpretability of the decision tree and the possibility of analyzing the results were among the advantages of this study.

Furthermore, Schamberg et al. (97) proposed a model for determining the appropriate dose of propofol using deep reinforcement learning based on the pain, sedation, and intensity (PSI) index. Sharma et al. (98) designed an optimal controller using type-2 fuzzy logic to determine the appropriate dose of sodium nitroprusside to maintain the patient's mean arterial pressure (MAP) at an appropriate level.

By reviewing the studies focusing on determining or predicting the appropriate dose of anesthetic drugs, it is evident that using deep recurrent neural networks, such as LSTM and GRU, and applying interpretable methods to explain the output of these networks lead to desirable outcomes, particularly when dealing with time series data.

Another critical issue in the discussion of anesthesia delivery control is the need for the physician to be informed of the drug concentration level in the patient's blood to determine the appropriate drug dose. Due to the challenges of performing complex calculations under operating conditions, devices have been designed and built to calculate the relevant drug concentration and report it to the physician. For instance, a study (99) used a model based on a support vector regression algorithm to compensate for errors in drug concentration measurements due to continuous sensor exposure to propofol, which might lead to sediment formation and inaccurate reporting of drug concentrations to the physician. Table 11 shows further details about each reviewed study in this category.

3.9. Category I: Monitoring

Monitoring involves the continuous assessment of a patient's hemodynamic status, including their cardiovascular and cerebral condition, in the operating room or the ICU, using specialized devices. Monitoring plays a crucial role in improving patient outcomes and the success of surgeries by maintaining vital signs within appropriate physiological ranges and quickly diagnosing and treating side effects before they lead to long-term complications. Therefore, designing highly accurate monitoring devices using AI and ML algorithms is necessary.

Artificial intelligence has been applied to monitor various aspects of a patient's condition. For example, blood pressure monitoring during anesthesia was evaluated using ML models, such as lasso restrictive logistic regression, neural networks, and SVM (100). The SVM model demonstrated superior performance based on the Kappa criterion, which measures the classifier's conformity in classifying samples.

False alarms from ICU vital signs monitors can be common, with rates ranging from 0.72 to 0.99. Machine learning models have been utilized to reduce the frequency of false alarms, and one study included missing sensor values in the input data of the model (101).

In the neurological critical care unit (NCCU), Unal et al. (102) used statistical techniques to determine the prevalence, types, and determinants of alarms. Liu et al. (103) conducted a study in 2020 on armpit temperature monitoring using an AI-enabled wireless, non-invasive armpit thermometer called iThermonitor. This thermometer provided accurate body temperature readings, compared to mercury thermometers. Additionally, a study (104) considered the validation of clinically relevant values of the relevant compensatory reserve measurement device with a dashboard view using a simple color code to diagnose bleeding. Table 12 shows key details of the reviewed studies in this category.

4. Discussion

This review study aimed to investigate the role of AI in anesthesia while also exploring the challenges, limitations, and opportunities in this field. It emphasizes the importance of having a realistic approach and appropriate expectations toward AI technologies in improving the treatment process.

Setting realistic expectations from AI techniques ensures that outcomes are well-defined and achievable, avoiding disappointment and vague results in their use. To facilitate an organized and coherent review, the selected studies were categorized into nine groups. Among these categories, event prediction and depth of anesthesia are the largest, with 43 and 21 articles, respectively.

The current review highlights the significant growth in the number of articles on event prediction in recent years, indicating the expanding research in this area. For example, predicting hypotension during anesthesia using ML models is a prominent topic in the perioperative subcategory. On the other hand, the depth of anesthesia category has a long history of research and offers various opportunities for further investigations, particularly in determining the depth of anesthesia using EEG signals and indices, such as BIS, through deep neural networks. However, certain categories, such as control of mechanical ventilation and weaning and operating room logistics, have limited studies and are considered emerging fields in AI applications for anesthesia.

One of the challenges revealed during the review is the difficulty of comparing studies in a similar field, mainly due to variations in research datasets and evaluation criteria. Privacy concerns in medical datasets often hinder accessibility and comparability. Researchers are encouraged to use publicly available datasets or release versions of the dataset while preserving individuals' privacy to facilitate further studies, result reproduction, and problem-solving. Moreover, using multiple evaluation criteria in research studies increases comparability and the validity of systematic reviews and meta-analyses.

The interpretability of ML and DL models is essential in informing the treatment team about prediction methods, boosting confidence in the models. For instance, in the ultrasound guidance category, where CNN is commonly used with image data, interpretability can be achieved through methods, such as Shapley and local interpretable model-agnostic explanations (LIME). Researchers should also employ feature ranking techniques to identify the most influential features in predictions for various ML models.

4.1. Conclusions

Using AI as a rapidly advancing technology can have a significant impact on various fields, including anesthesiology. This review highlights the role of AI models in establishing monitoring and decision support systems in the domain of anesthesia. The studies reviewed were categorized into nine distinct areas, and the materials in each study and category were presented in an organized and tabular manner. Each section also included suggestions for future work and ideas.

The continuous progress in AI techniques offers great potential to support anesthesiologists in enhancing their performance. By carefully and judiciously employing AI approaches, it is possible to improve anesthesia-related tasks and patient care. Additionally, given the importance of interpretability in medical decision-making, using interpretable AI techniques is strongly recommended for future studies. These methods allow physicians to analyze and understand the results, leading to more confident and informed

Table 12. Category I: Monitoring

No.	Study	Goal	Type of Surgery	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance
1	Pasma et al. (2021) (100)	Automated artifact removal in anesthesia blood pressure data	Non-cardiac and non-thoracic	Available	88(M:39,F:49)/University Medical Center Utrecht	Feature Types: Systolic blood pressure, diastolic blood pressure, mean blood pressure, heart rate, pulse pressure (systolic–diastolic blood pressure), ratios between heartrate and blood pressure (systolic blood pressure divided by heartrate and diastolic blood pressure divided by heartrate), ratio between systolic and mean arterial blood pressure, and ratio between mean and diastolic arterial blood pressure	Lasso Restrictive LR, NN, SVM	SVM	Kappa: from 0.524 to 0.651
2	Hever et al. (2020) (101)	To analyze in real-time missing sensor data to minimize false alarm rate	-	Available	481(M:325,F:156)/Shanghai Jiao Tong University School of Medicine affiliated Ruijin Hospital	Age, Gender, NC, WC, BMI and faciocervical measurements (maximum interincisal distance (MID)), height to thyrosternum distance (H/TSD))	SABIHC2 (It is a machine learning model based on SVM) & STOP-BANG (It is one of the most widely used questionnaires)	SABIHC2	AUC: 0.832; Sensitivity: 91.6%; Specificity: 74.9%

Abbreviations: LR, logistic regression; NN, neural network; SVM, support vector machine; NC, neck circumference; WC, waist circumference; BMI, body mass index; AUC, area under the curve.

decisions. Overall, the integration of AI in anesthesia holds promise for optimizing patient outcomes and the overall efficiency of anesthesia management. As technology advances, researchers and practitioners should continue exploring innovative AI applications to further revolutionize the field of anesthesiology.

Footnotes

Authors' Contribution: Sana Hashemi: Collecting the articles, drafting of the manuscript; Zohreh Yousefzadeh: Collecting the articles, drafting of the manuscript; Ahmad Ali Abin: Study concept and design, study supervision, drafting of the manuscript; Azar Ejmalian: Study concept and design, collecting and categorizing the articles, study supervision, critical revision of the manuscript; Shahabedin Nabavi: Critical revision of the manuscript; Ali Dabbagh: Critical revision of the manuscript.

Conflict of Interests: There is no conflict of interest to declare.

Funding/Support: This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Nabavi S, Ejmalian A, Moghaddam ME, Abin AA, Frangi AF, Mohammadi M, et al. Medical imaging and computational image analysis in COVID-19 diagnosis: A review. *Comput Biol Med*. 2021;**135**:104605. [PubMed ID: 34175533]. [PubMed Central ID: PMC8219713]. https://doi.org/10.1016/j.compbiomed.2021.104605.
2. Badrigilan S, Nabavi S, Abin AA, Rostampour N, Abedi I, Shirvani A, et al. Deep learning approaches for automated classification and segmentation of head and neck cancers and brain tumors in magnetic resonance images: A meta-analysis study. *Int J Comput Assist Radiol Surg*. 2021;**16**(4):529-42. [PubMed ID: 33666859]. https://doi.org/10.1007/s11548-021-02326-z.
3. Madanu R, Rahman F, Abbod MF, Fan SZ, Shieh JS. Depth of anesthesia prediction via EEG signals using convolutional neural network and ensemble empirical mode decomposition. *Math Biosci Eng*. 2021;**18**(5):5047-68. [PubMed ID: 34517477]. https://doi.org/10.3934/mbe.2021257.
4. Schweingruber N, Mader MMD, Wiehe A, Roder F, Gottsche J, Kluge S, et al. A recurrent machine learning model predicts intracranial hypertension in neurointensive care patients. *Brain*. 2022;**145**(8):2910-9. [PubMed ID: 35139181]. [PubMed Central ID: PMC9486888]. https://doi.org/10.1093/brain/awab453.
5. Bishara A, Chiu C, Whitlock EL, Douglas VC, Lee S, Butte AJ, et al. Postoperative delirium prediction using machine learning models and preoperative electronic health record data. *BMC Anesthesiol*. 2022;**22**(1):8. [PubMed ID: 34979919]. [PubMed Central ID: PMC8722098]. https://doi.org/10.1186/s12871-021-01543-y.
6. Hashimoto DA, Witkowski E, Gao L, Meireles O, Rosman G. Artificial intelligence in anesthesiology: Current techniques, clinical applications, and limitations. *Anesthesiology*. 2020;**132**(2):379-94. [PubMed ID: 31939856]. [PubMed Central ID: PMC7643051]. https://doi.org/10.1097/ALN.0000000000002960.
7. Abin AA. Querying beneficial constraints before clustering using facility location analysis. *IEEE Trans Cybern*. 2018;**48**(1):312-23. [PubMed ID: 27959838]. https://doi.org/10.1109/TCYB.2016.2633498.
8. Ejmalian A, Aghaei A, Nabavi S, Abedzadeh Darabad M, Tajbakhsh A, Abin AA, et al. Prediction of acute kidney injury after cardiac surgery using interpretable machine learning. *Anesth Pain Med*. 2022;**12**(4). e127140. [PubMed ID: 36937087]. [PubMed Central ID: PMC10016126]. https://doi.org/10.5812/aapm-127140.

9. Xin D, An Z, Ding J, Li Z, Qiao L. Magnetic resonance imaging features under deep learning algorithms in evaluated cerebral protection of craniotomy evacuation of hematoma under propofol anesthesia. *Contrast Media Mol Imaging*. 2021;**2021**:2209527. [PubMed ID: 34671228]. [PubMed Central ID: PMC8500760]. https://doi.org/10.1155/2021/2209527.

10. Narula G, Haeberlin M, Balsiger J, Strassle C, Imbach LL, Keller E. Detection of EEG burst-suppression in neurocritical care patients using an unsupervised machine learning algorithm. *Clin Neurophysiol*. 2021;**132**(10):2485-92. [PubMed ID: 34454277]. https://doi.org/10.1016/j.clinph.2021.07.018.

11. Farzaneh N, Williamson CA, Jiang C, Srinivasan A, Bapuraj JR, Gryak J, et al. Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries. *Diagnostics (Basel)*. 2020;**10**(10). [PubMed ID: 33007929]. [PubMed Central ID: PMC7600198]. https://doi.org/10.3390/diagnostics10100773.

12. Farzaneh N, Williamson CA, Gryak J, Najarian K. A hierarchical expert-guided machine learning framework for clinical decision support systems: An application to traumatic brain injury prognostication. *NPJ Digit Med*. 2021;**4**(1):78. [PubMed ID: 33963275]. [PubMed Central ID: PMC8105342]. https://doi.org/10.1038/s41746-021-00445-0.

13. Bernabei JM, Owoputi O, Small SD, Nyema NT, Dumenyo E, Kim J, et al. A full-stack application for detecting seizures and reducing data during continuous electroencephalogram monitoring. *Crit Care Explor*. 2021;**3**(7). e0476. [PubMed ID: 34278312]. [PubMed Central ID: PMC8280012]. https://doi.org/10.1097/CCE.0000000000000476.

14. Fumeaux NF, Ebrahim S, Coughlin BF, Kadambi A, Azmi A, Xu JX, et al. Accurate detection of spontaneous seizures using a generalized linear model with external validation. *Epilepsia*. 2020;**61**(9):1906-18. [PubMed ID: 32761902]. [PubMed Central ID: PMC7953845]. https://doi.org/10.1111/epi.16628.

15. Koch M, Acharjee A, Ament Z, Schleicher R, Bevers M, Stapleton C, et al. Machine learning-driven metabolomic evaluation of cerebrospinal fluid: Insights into poor outcomes after aneurysmal subarachnoid hemorrhage. *Neurosurgery*. 2021;**88**(5):1003-11. [PubMed ID: 33469656]. [PubMed Central ID: PMC8046589]. https://doi.org/10.1093/neuros/nyaa557.

16. Elmer J, Coppler PJ, May TL, Hirsch K, Faro J, Solanki P, et al. Unsupervised learning of early post-arrest brain injury phenotypes. *Resuscitation*. 2020;**153**:154-60. [PubMed ID: 32531403]. [PubMed Central ID: PMC7390683]. https://doi.org/10.1016/j.resuscitation.2020.05.051.

17. Tan HS, Liu N, Sultana R, Han NR, Tan CW, Zhang J, et al. Prediction of breakthrough pain during labour neuraxial analgesia: Comparison of machine learning and multivariable regression approaches. *Int J Obstet Anesth*. 2021;**45**:99-110. [PubMed ID: 33121883]. https://doi.org/10.1016/j.ijoa.2020.08.010.

18. Gonzalez-Cava JM, Arnay R, Leon A, Martin M, Reboso JA, Calvo-Rolle JL, et al. Machine learning based method for the evaluation of the Analgesia Nociception Index in the assessment of general anesthesia. *Comput Biol Med*. 2020;**118**:103645. [PubMed ID: 32174322]. https://doi.org/10.1016/j.compbiomed.2020.103645.

19. Meijer F, Honing M, Roor T, Toet S, Calis P, Olofsen E, et al. Reduced postoperative pain using Nociception Level-guided fentanyl dosing during sevoflurane anaesthesia: A randomised controlled trial. *Br J Anaesth*. 2020;**125**(6):1070-8. [PubMed ID: 32950246]. [PubMed Central ID: PMC7771114]. https://doi.org/10.1016/j.bja.2020.07.057.

20. Choi BM, Yim JY, Shin H, Noh G. Novel analgesic index for postoperative pain assessment based on a photoplethysmographic spectrogram and convolutional neural network: Observational study. *J Med Internet Res*. 2021;**23**(2). e23920. [PubMed ID: 33533723]. [PubMed Central ID: PMC7889419]. https://doi.org/10.2196/23920.

21. Barry GS, Bailey JG, Sardinha J, Brousseau P, Uppal V. Factors associated with rebound pain after peripheral nerve block for ambulatory surgery. *Br J Anaesth*. 2021;**126**(4):862-71. [PubMed ID: 33390261]. https://doi.org/10.1016/j.bja.2020.10.035.

22. Teichmann D, Hallmann A, Wolfart S, Teichmann M. Identification of dental pain sensation based on cardiorespiratory signals. *Biomed Tech (Berl)*. 2021;**66**(2):159-65. [PubMed ID: 33768763]. https://doi.org/10.1515/bmt-2020-0047.

23. Zhang Z, Liu J, Xi J, Gong Y, Zeng L, Ma P. Derivation and validation of an ensemble model for the prediction of agitation in mechanically ventilated patients maintained under light sedation. *Crit Care Med*. 2021;**49**(3):e279-90. [PubMed ID: 33470778]. https://doi.org/10.1097/CCM.0000000000004821.

24. Ang CYS, Chiew YS, Vu LH, Cove ME. Quantification of respiratory effort magnitude in spontaneous breathing patients using convolutional autoencoders. *Comput Methods Programs Biomed*. 2022;**215**:106601. [PubMed ID: 34973606]. https://doi.org/10.1016/j.cmpb.2021.106601.

25. Solomon SC, Saxena RC, Neradilek MB, Hau V, Fong CT, Lang JD, et al. Forecasting a Crisis: Machine-learning models predict occurrence of intraoperative bradycardia associated with hypotension. *Anesth Analg*. 2020;**130**(5):1201-10. [PubMed ID: 32287127]. https://doi.org/10.1213/ANE.0000000000004636.

26. Jalali A, Lonsdale H, Zamora LV, Ahumada L, Nguyen ATH, Rehman M, et al. Machine learning applied to registry data: Development of a patient-specific prediction model for blood transfusion requirements during craniofacial surgery using the pediatric craniofacial perioperative registry dataset. *Anesth Analg*. 2021;**132**(1):160-71. [PubMed ID: 32618624]. https://doi.org/10.1213/ANE.0000000000004988.

27. Maheshwari K, Buddi S, Jian Z, Settels J, Shimada T, Cohen B, et al. Performance of the Hypotension Prediction Index with non-invasive arterial pressure waveforms in non-cardiac surgical patients. *J Clin Monit Comput*. 2021;**35**(1):71-8. [PubMed ID: 31989416]. [PubMed Central ID: PMC7889685]. https://doi.org/10.1007/s10877-020-00463-5.

28. Li XF, Huang YZ, Tang JY, Li RC, Wang XQ. Development of a random forest model for hypotension prediction after anesthesia induction for cardiac surgery. *World J Clin Cases*. 2021;**9**(29):8729-39. [PubMed ID: 34734051]. [PubMed Central ID: PMC8546817]. https://doi.org/10.12998/wjcc.v9.i29.8729.

29. Frassanito L, Giuri PP, Vassalli F, Piersanti A, Longo A, Zanfini BA, et al. Hypotension Prediction Index with non-invasive continuous arterial pressure waveforms (ClearSight): clinical performance in Gynaecologic Oncologic Surgery. *J Clin Monit Comput*. 2022;**36**(5):1325-32. [PubMed ID: 34618291]. [PubMed Central ID: PMC8496438]. https://doi.org/10.1007/s10877-021-00763-4.

30. Gratz I, Baruch M, Takla M, Seaman J, Allen I, McEniry B, et al. The application of a neural network to predict hypotension and vasopressor requirements non-invasively in obstetric patients having spinal anesthesia for elective cesarean section (C/S). *BMC Anesthesiol*. 2020;**20**(1):98. [PubMed ID: 32357833]. [PubMed Central ID: PMC7195764]. https://doi.org/10.1186/s12871-020-01015-9.

31. Lee J, Woo J, Kang AR, Jeong YS, Jung W, Lee M, et al. Comparative Analysis on Machine Learning and Deep Learning to Predict Post-Induction Hypotension. *Sensors (Basel)*. 2020;**20**(16). [PubMed ID: 32824073]. [PubMed Central ID: PMC7472016]. https://doi.org/10.3390/s20164575.

32. Kang AR, Lee J, Jung W, Lee M, Park SY, Woo J, et al. Development of a prediction model for hypotension after induction of anesthesia using machine learning. *PLoS One*. 2020;**15**(4). e0231172. [PubMed ID: 32298292]. [PubMed Central ID: PMC7162491]. https://doi.org/10.1371/journal.pone.0231172.

33. Wijnberge M, Geerts BF, Hol L, Lemmers N, Mulder MP, Berge P, et al. Effect of a Machine Learning-Derived Early Warning System for Intraoperative Hypotension vs Standard Care on Depth and Duration of Intraoperative Hypotension During Elective Noncardiac Surgery:

The HYPE Randomized Clinical Trial. *JAMA*. 2020;**323**(11):1052-60. [PubMed ID: 32065827]. [PubMed Central ID: PMC7078808]. https://doi.org/10.1001/jama.2020.0592.

34. Kim JH, Choi JW, Kwon YS, Kang SS. Predictive model for difficult laryngoscopy using machine learning: retrospective cohort study. *Braz J Anesthesiol*. 2022;**72**(5):622-8. [PubMed ID: 34252452]. [PubMed Central ID: PMC9515663]. https://doi.org/10.1016/j.bjane.2021.06.016.
35. Kim JH, Kim H, Jang JS, Hwang SM, Lim SY, Lee JJ, et al. Development and validation of a difficult laryngoscopy prediction model using machine learning of neck circumference and thyromental height. *BMC Anesthesiol*. 2021;**21**(1):125. [PubMed ID: 33882838]. [PubMed Central ID: PMC8059322]. https://doi.org/10.1186/s12871-021-01343-4.
36. Mathis MR, Engoren MC, Joo H, Maile MD, Aaronson KD, Burns ML, et al. Early Detection of Heart Failure With Reduced Ejection Fraction Using Perioperative Data Among Noncardiac Surgical Patients: A Machine-Learning Approach. *Anesth Analg*. 2020;**130**(5):1188-200. [PubMed ID: 32287126]. [PubMed Central ID: PMC7467779]. https://doi.org/10.1213/ANE.0000000000004630.
37. Bollepalli SC, Sevakula RK, Au-Yeung WM, Kassab MB, Merchant FM, Bazoukis G, et al. Real-Time Arrhythmia Detection Using Hybrid Convolutional Neural Networks. *J Am Heart Assoc*. 2021;**10**(23). e023222. [PubMed ID: 34854319]. [PubMed Central ID: PMC9075394]. https://doi.org/10.1161/JAHA.121.023222.
38. Yeh LR, Chen WC, Chan HY, Lu NH, Wang CY, Twan WH, et al. Integrating ECG Monitoring and Classification via IoT and Deep Neural Networks. *Biosensors (Basel)*. 2021;**11**(6). [PubMed ID: 34201215]. [PubMed Central ID: PMC8226863]. https://doi.org/10.3390/bios11060188.
39. Hadjipavlou G, Warnaby CE, Fitzgerald J, Sleigh J. Contributions of synaptic and astrocyte physiology to the anaesthetised encephalogram revealed using a computational model. *Br J Anaesth*. 2021;**126**(5):985-95. [PubMed ID: 33773753]. [PubMed Central ID: PMC8132884]. https://doi.org/10.1016/j.bja.2021.01.034.
40. Lee CK, Samad M, Hofer I, Cannesson M, Baldi P. Development and validation of an interpretable neural network for prediction of postoperative in-hospital mortality. *NPJ Digit Med*. 2021;**4**(1):8. [PubMed ID: 33420341]. [PubMed Central ID: PMC7794438]. https://doi.org/10.1038/s41746-020-00377-1.
41. Chelazzi C, Villa G, Manno A, Ranfagni V, Gemmi E, Romagnoli S. The new SUMPOT to predict postoperative complications using an Artificial Neural Network. *Sci Rep*. 2021;**11**(1):22692. [PubMed ID: 34811383]. [PubMed Central ID: PMC8608915]. https://doi.org/10.1038/s41598-021-01913-z.
42. Racine AM, Tommet D, D'Aquila ML, Fong TG, Gou Y, Tabloski PA, et al. Machine Learning to Develop and Internally Validate a Predictive Model for Post-operative Delirium in a Prospective, Observational Clinical Cohort Study of Older Surgical Patients. *J Gen Intern Med*. 2021;**36**(2):265-73. [PubMed ID: 33078300]. [PubMed Central ID: PMC7878663]. https://doi.org/10.1007/s11606-020-06238-7.
43. Palla K, Hyland SL, Posner K, Ghosh P, Nair B, Bristow M, et al. Intraoperative prediction of postanaesthesia care unit hypotension. *Br J Anaesth*. 2022;**128**(4):623-35. [PubMed ID: 34924175]. [PubMed Central ID: PMC9074793]. https://doi.org/10.1016/j.bja.2021.10.052.
44. Schenk J, Wijnberge M, Maaskant JM, Hollmann MW, Hol L, Immink RV, et al. Effect of Hypotension Prediction Index-guided intraoperative haemodynamic care on depth and duration of postoperative hypotension: a sub-study of the Hypotension Prediction trial. *Br J Anaesth*. 2021;**127**(5):681-8. [PubMed ID: 34303491]. https://doi.org/10.1016/j.bja.2021.05.033.
45. Tan J, Wang Q, Shi W, Liang K, Yu B, Mao Q. A Machine Learning Approach for Predicting Early Phase Postoperative Hypertension in Patients Undergoing Carotid Endarterectomy. *Ann Vasc Surg*. 2021;**71**:121-31. [PubMed ID: 32653616]. https://doi.org/10.1016/j.avsg.2020.07.001.
46. Jeong YS, Kim J, Kim D, Woo J, Kim MG, Choi HW, et al. Prediction of Postoperative Complications for Patients of End Stage Renal Disease. *Sensors (Basel)*. 2021;**21**(2). [PubMed ID: 33466610]. [PubMed Central ID: PMC7828737]. https://doi.org/10.3390/s21020544.
47. Bai P, Zhou Y, Liu Y, Li G, Li Z, Wang T, et al. Risk Factors of Cerebral Infarction and Myocardial Infarction after Carotid Endarterectomy Analyzed by Machine Learning. *Comput Math Methods Med*. 2020;**2020**:6217392. [PubMed ID: 33273961]. [PubMed Central ID: PMC7683166]. https://doi.org/10.1155/2020/6217392.
48. Ko S, Jo C, Chang CB, Lee YS, Moon YW, Youm JW, et al. A web-based machine-learning algorithm predicting postoperative acute kidney injury after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc*. 2022;**30**(2):545-54. [PubMed ID: 32880677]. https://doi.org/10.1007/s00167-020-06258-0.
49. Cao Y, Montgomery S, Ottosson J, Naslund E, Stenberg E. Deep Learning Neural Networks to Predict Serious Complications After Bariatric Surgery: Analysis of Scandinavian Obesity Surgery Registry Data. *JMIR Med Inform*. 2020;**8**(5). e15992. [PubMed ID: 32383681]. [PubMed Central ID: PMC7244994]. https://doi.org/10.2196/15992.
50. Qian K, Wu S, Lee W, Liu S, Li A, Cang J, et al. A model-based validation study of postoperative complications with considerations on operative timing. *Ann Transl Med*. 2021;**9**(8):708. [PubMed ID: 33987406]. [PubMed Central ID: PMC8106084]. https://doi.org/10.21037/atm-21-669.
51. Suhre W, O'Reilly-Shah V, Van Cleve W. Cannabis use is associated with a small increase in the risk of postoperative nausea and vomiting: a retrospective machine-learning causal analysis. *BMC Anesthesiol*. 2020;**20**(1):115. [PubMed ID: 32423445]. [PubMed Central ID: PMC7236204]. https://doi.org/10.1186/s12871-020-01036-4.
52. Lu Y, Forlenza E, Cohn MR, Lavoie-Gagne O, Wilbur RR, Song BM, et al. Machine learning can reliably identify patients at risk of overnight hospital admission following anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc*. 2021;**29**(9):2958-66. [PubMed ID: 33047150]. https://doi.org/10.1007/s00167-020-06321-w.
53. Lu Y, Kunze K, Cohn MR, Lavoie-Gagne O, Polce E, Nwachukwu BU, et al. Artificial Intelligence Predicts Cost After Ambulatory Anterior Cruciate Ligament Reconstruction. *Arthrosc Sports Med Rehabil*. 2021;**3**(6):e2033-45. [PubMed ID: 34977663]. [PubMed Central ID: PMC8689347]. https://doi.org/10.1016/j.asmr.2021.10.013.
54. Magunia H, Lederer S, Verbuecheln R, Gilot BJ, Koeppen M, Haeberle HA, et al. Machine learning identifies ICU outcome predictors in a multicenter COVID-19 cohort. *Crit Care*. 2021;**25**(1):295. [PubMed ID: 34404458]. [PubMed Central ID: PMC8370055]. https://doi.org/10.1186/s13054-021-03720-4.
55. Yun K, Oh J, Hong TH, Kim EY. Prediction of Mortality in Surgical Intensive Care Unit Patients Using Machine Learning Algorithms. *Front Med (Lausanne)*. 2021;**8**:621861. [PubMed ID: 33869245]. [PubMed Central ID: PMC8044535]. https://doi.org/10.3389/fmed.2021.621861.
56. Cherifa M, Interian Y, Blet A, Resche-Rigon M, Pirracchio R. The Physiological Deep Learner: First application of multitask deep learning to predict hypotension in critically ill patients. *Artif Intell Med*. 2021;**118**:102118. [PubMed ID: 34412841]. https://doi.org/10.1016/j.artmed.2021.102118.
57. Moghadam MC, Abad EMK, Bagherzadeh N, Ramsingh D, Li GP, Kain ZN. A machine-learning approach to predicting hypotensive events in ICU settings. *Comput Biol Med*. 2020;**118**:103626. [PubMed ID: 32174328]. https://doi.org/10.1016/j.compbiomed.2020.103626.
58. Cherifa M, Blet A, Chambaz A, Gayat E, Resche-Rigon M, Pirracchio R. Prediction of an Acute Hypotensive Episode During an ICU Hospitalization With a Super Learner Machine-Learning Algorithm. *Anesth Analg*. 2020;**130**(5):1157-66. [PubMed ID: 32287123]. https://doi.org/10.1213/ANE.0000000000004539.

59. Hu J, Fung FW, Jacobwitz M, Parikh DS, Vala L, Donnelly M, et al. Machine learning models to predict electroencephalographic seizures in critically ill children. *Seizure*. 2021;**87**:61-8. [PubMed ID: 33714840]. [PubMed Central ID: PMC8044039]. https://doi.org/10.1016/j.seizure.2021.03.001.

60. Chang CC, Yeh JH, Chiu HC, Chen YM, Jhou MJ, Liu TC, et al. Utilization of Decision Tree Algorithms for Supporting the Prediction of Intensive Care Unit Admission of Myasthenia Gravis: A Machine Learning-Based Approach. *J Pers Med*. 2022;**12**(1). [PubMed ID: 35055347]. [PubMed Central ID: PMC8778268]. https://doi.org/10.3390/jpm12010032.

61. Hayasaka T, Kawano K, Kurihara K, Suzuki H, Nakane M, Kawamae K. Creation of an artificial intelligence model for intubation difficulty classification by deep learning (convolutional neural network) using face images: an observational study. *J Intensive Care*. 2021;**9**(1):38. [PubMed ID: 33952341]. [PubMed Central ID: PMC8101256]. https://doi.org/10.1186/s40560-021-00551-x.

62. Wu YC, Wong LT, Wu CL, Chao WC. The association between culture positivity and long-term mortality in critically ill surgical patients. *J Intensive Care*. 2021;**9**(1):66. [PubMed ID: 34702345]. [PubMed Central ID: PMC8546784]. https://doi.org/10.1186/s40560-021-00576-2.

63. Ling J, Liao T, Wu Y, Wang Z, Jin H, Lu F, et al. Predictive value of red blood cell distribution width in septic shock patients with thrombocytopenia: A retrospective study using machine learning. *J Clin Lab Anal*. 2021;**35**(12). e24053. [PubMed ID: 34674393]. [PubMed Central ID: PMC8649348]. https://doi.org/10.1002/jcla.24053.

64. Sarton B, Jaquet P, Belkacemi D, de Montmollin E, Bonneville F, Sazio C, et al. Assessment of Magnetic Resonance Imaging Changes and Functional Outcomes Among Adults With Severe Herpes Simplex Encephalitis. *JAMA Netw Open*. 2021;**4**(7). e2114328. [PubMed ID: 34313743]. [PubMed Central ID: PMC8317014]. https://doi.org/10.1001/jamanetworkopen.2021.14328.

65. Elmer J, Jones BL, Nagin DS. Comparison of parametric and nonparametric methods for outcome prediction using longitudinal data after cardiac arrest. *Resuscitation*. 2020;**148**:152-60. [PubMed ID: 32004661]. [PubMed Central ID: PMC7132134]. https://doi.org/10.1016/j.resuscitation.2020.01.020.

66. Mwikirize C, Kimbowa AB, Imanirakiza S, Katumba A, Nosher JL, Hacihaliloglu I. Time-aware deep neural networks for needle tip localization in 2D ultrasound. *Int J Comput Assist Radiol Surg*. 2021;**16**(5):819-27. [PubMed ID: 33840037]. https://doi.org/10.1007/s11548-021-02361-w.

67. Liu Y, Cheng L. Ultrasound Images Guided under Deep Learning in the Anesthesia Effect of the Regional Nerve Block on Scapular Fracture Surgery. *J Healthc Eng*. 2021;**2021**:6231116. [PubMed ID: 34659690]. [PubMed Central ID: PMC8516573]. https://doi.org/10.1155/2021/6231116.

68. Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;**39**(12):2481-95. [PubMed ID: 28060704]. https://doi.org/10.1109/TPAMI.2016.2644615.

69. Bowness J, Varsou O, Turbitt L, Burkett-St Laurent D. Identifying anatomical structures on ultrasound: assistive artificial intelligence in ultrasound-guided regional anesthesia. *Clin Anat*. 2021;**34**(5):802-9. [PubMed ID: 33904628]. https://doi.org/10.1002/ca.23742.

70. Paris A, Hafiane A. Shape constraint function for artery tracking in ultrasound images. *Comput Med Imaging Graph*. 2021;**93**:101970. [PubMed ID: 34428649]. https://doi.org/10.1016/j.compmedimag.2021.101970.

71. Gabriel RA, Harjai B, Simpson S, Goldhaber N, Curran BP, Waterman RS. Machine Learning-Based Models Predicting Outpatient Surgery End Time and Recovery Room Discharge at an Ambulatory Surgery Center. *Anesth Analg*. 2022;**135**(1):159-69. [PubMed ID: 35389380]. [PubMed Central ID: PMC9172889]. https://doi.org/10.1213/ANE.0000000000006015.

72. Jiao Y, Sharma A, Ben Abdallah A, Maddox TM, Kannampallil T. Probabilistic forecasting of surgical case duration using machine learning: model development and validation. *J Am Med Inform Assoc*. 2020;**27**(12):1885-93. [PubMed ID: 33031543]. [PubMed Central ID: PMC7727362]. https://doi.org/10.1093/jamia/ocaa140.

73. Liu L, Ni Y, Beck AF, Brokamp C, Ramphul RC, Highfield LD, et al. Understanding Pediatric Surgery Cancellation: Geospatial Analysis. *J Med Internet Res*. 2021;**23**(9). e26231. [PubMed ID: 34505837]. [PubMed Central ID: PMC8463951]. https://doi.org/10.2196/26231.

74. Abel JH, Badgeley MA, Meschede-Krasa B, Schamberg G, Garwood IC, Lecamwasam K, et al. Machine learning of EEG spectra classifies unconsciousness during GABAergic anesthesia. *PLoS One*. 2021;**16**(5). e0246165. [PubMed ID: 33956800]. [PubMed Central ID: PMC8101756]. https://doi.org/10.1371/journal.pone.0246165.

75. Afshar S, Boostani R, Sanei S. A Combinatorial Deep Learning Structure for Precise Depth of Anesthesia Estimation From EEG Signals. *IEEE J Biomed Health Inform*. 2021;**25**(9):3408-15. [PubMed ID: 33760743]. https://doi.org/10.1109/JBHI.2021.3068481.

76. Duclos C, Maschke C, Mahdid Y, Berkun K, Castanheira JDS, Tarnal V, et al. Differential classification of states of consciousness using envelope- and phase-based functional connectivity. *Neuroimage*. 2021;**237**:118171. [PubMed ID: 34000405]. https://doi.org/10.1016/j.neuroimage.2021.118171.

77. Avilov O, Rimbert S, Popov A, Bougrain L. Optimizing Motor Intention Detection With Deep Learning: Towards Management of Intraoperative Awareness. *IEEE Trans Biomed Eng*. 2021;**68**(10):3087-97. [PubMed ID: 33687833]. https://doi.org/10.1109/TBME.2021.3064794.

78. Ra JS, Li T, Li Y. A novel spectral entropy-based index for assessing the depth of anaesthesia. *Brain Inform*. 2021;**8**(1):10. [PubMed ID: 33978842]. [PubMed Central ID: PMC8116386]. https://doi.org/10.1186/s40708-021-00130-8.

79. Li F, Li Y, Zheng H, Jiang L, Gao D, Li C, et al. Identification of the General Anesthesia Induced Loss of Consciousness by Cross Fuzzy Entropy-Based Brain Network. *IEEE Trans Neural Syst Rehabil Eng*. 2021;**29**:2281-91. [PubMed ID: 34705652]. https://doi.org/10.1109/TNSRE.2021.3123696.

80. Li Y, Li F, Zheng H, Jiang L, Peng Y, Zhang Y, et al. Recognition of general anesthesia-induced loss of consciousness based on the spatial pattern of the brain networks. *J Neural Eng*. 2021;**18**(5). [PubMed ID: 34534980]. https://doi.org/10.1088/1741-2552/ac27fc.

81. Ramaswamy SM, Kuizenga MH, Weerink MAS, Vereecke HEM, Struys M, Belur Nagaraj S. Frontal electroencephalogram based drug, sex, and age independent sedation level prediction using non-linear machine learning algorithms. *J Clin Monit Comput*. 2022;**36**(1):121-30. [PubMed ID: 33315176]. [PubMed Central ID: PMC7734899]. https://doi.org/10.1007/s10877-020-00627-3.

82. Kashkooli K, Polk SL, Hahm EY, Murphy J, Ethridge BR, Gitlin J, et al. Improved tracking of sevoflurane anesthetic states with drug-specific machine learning models. *J Neural Eng*. 2020;**17**(4):46020. [PubMed ID: 32485685]. [PubMed Central ID: PMC7540939]. https://doi.org/10.1088/1741-2552/ab98da.

83. Lee H, Wang S, Hudetz AG. State-Dependent Cortical Unit Activity Reflects Dynamic Brain State Transitions in Anesthesia. *J Neurosci*. 2020;**40**(49):9440-54. [PubMed ID: 33122389]. [PubMed Central ID: PMC7724136]. https://doi.org/10.1523/JNEUROSCI.0601-20.2020.

84. Hayase K, Hayashi K, Sawa T. Hierarchical Poincare analysis for anaesthesia monitoring. *J Clin Monit Comput*. 2020;**34**(6):1321-30. [PubMed ID: 31863245]. https://doi.org/10.1007/s10877-019-00447-0.

85. Shalbaf A, Shalbaf R, Saffar M, Sleigh J. Monitoring the level of hypnosis using a hierarchical SVM system. *J Clin Monit Comput*.

2020;**34**(2):331-8. [PubMed ID: 30982945]. https://doi.org/10.1007/s10877-019-00311-1.

86. Park Y, Han SH, Byun W, Kim JH, Lee HC, Kim SJ. A Real-Time Depth of Anesthesia Monitoring System Based on Deep Neural Network With Large EDO Tolerant EEG Analog Front-End. *IEEE Trans Biomed Circuits Syst*. 2020;**14**(4):825-37. [PubMed ID: 32746339]. https://doi.org/10.1109/TBCAS.2020.2998172.

87. Li R, Wu Q, Liu J, Wu Q, Li C, Zhao Q. Monitoring Depth of Anesthesia Based on Hybrid Features and Recurrent Neural Network. *Front Neurosci*. 2020;**14**:26. [PubMed ID: 32116494]. [PubMed Central ID: PMC7020827]. https://doi.org/10.3389/fnins.2020.00026.

88. Ihalainen R, Gosseries O, de Steen FV, Raimondo F, Panda R, Bonhomme V, et al. How hot is the hot zone? Computational modelling clarifies the role of parietal and frontoparietal connectivity during anaesthetic-induced loss of consciousness. *Neuroimage*. 2021;**231**:117841. [PubMed ID: 33577934]. https://doi.org/10.1016/j.neuroimage.2021.117841.

89. Dubost C, Humbert P, Oudre L, Labourdette C, Vayatis N, Vidal PP. Quantitative assessment of consciousness during anesthesia without EEG data. *J Clin Monit Comput*. 2021;**35**(5):993-1005. [PubMed ID: 32661827]. https://doi.org/10.1007/s10877-020-00553-4.

90. Zhan J, Wu ZX, Duan ZX, Yang GY, Du ZY, Bao XH, et al. Heart rate variability-derived features based on deep neural network for distinguishing different anaesthesia states. *BMC Anesthesiol*. 2021;**21**(1):66. [PubMed ID: 33653263]. [PubMed Central ID: PMC7923817]. https://doi.org/10.1186/s12871-021-01285-x.

91. Campbell JM, Huang Z, Zhang J, Wu X, Qin P, Northoff G, et al. Pharmacologically informed machine learning approach for identifying pathological states of unconsciousness via resting-state fMRI. *Neuroimage*. 2020;**206**:116316. [PubMed ID: 31672663]. [PubMed Central ID: PMC6981054]. https://doi.org/10.1016/j.neuroimage.2019.116316.

92. Sanz Perl Y, Pallavicini C, Perez Ipina I, Demertzi A, Bonhomme V, Martial C, et al. Perturbations in dynamical models of whole-brain activity dissociate between the level and stability of consciousness. *PLoS Comput Biol*. 2021;**17**(7). e1009139. [PubMed ID: 34314430]. [PubMed Central ID: PMC8315553]. https://doi.org/10.1371/journal.pcbi.1009139.

93. Tacke M, Kochs EF, Mueller M, Kramer S, Jordan D, Schneider G. Machine learning for a combined electroencephalographic anesthesia index to detect awareness under anesthesia. *PLoS One*. 2020;**15**(8). e0238249. [PubMed ID: 32845935]. [PubMed Central ID: PMC7449480]. https://doi.org/10.1371/journal.pone.0238249.

94. Ingrande J, Gabriel RA, McAuley J, Krasinska K, Chien A, Lemmens HJM. The Performance of an Artificial Neural Network Model in Predicting the Early Distribution Kinetics of Propofol in Morbidly Obese and Lean Subjects. *Anesth Analg*. 2020;**131**(5):1500-9. [PubMed ID: 33079873]. https://doi.org/10.1213/ANE.0000000000004897.

95. Miyaguchi N, Takeuchi K, Kashima H, Morita M, Morimatsu H. Predicting anesthetic infusion events using machine learning. *Sci Rep*. 2021;**11**(1):23648. [PubMed ID: 34880365]. [PubMed Central ID: PMC8655034]. https://doi.org/10.1038/s41598-021-03112-2.

96. Wei CN, Wang LY, Chang XY, Zhou QH. A prediction model using machine-learning algorithm for assessing intrathecal hyperbaric bupivacaine dose during cesarean section. *BMC Anesthesiol*. 2021;**21**(1):116. [PubMed ID: 33853548]. [PubMed Central ID: PMC8045295]. https://doi.org/10.1186/s12871-021-01331-8.

97. Schamberg G, Badgeley M, Meschede-Krasa B, Kwon O, Brown EN. Continuous action deep reinforcement learning for propofol dosing during general anesthesia. *Artif Intell Med*. 2022;**123**:102227. [PubMed ID: 34998516]. https://doi.org/10.1016/j.artmed.2021.102227.

98. Sharma R, Deepak KK, Gaur P, Joshi D. An optimal interval type-2 fuzzy logic control based closed-loop drug administration to regulate the mean arterial blood pressure. *Comput Methods Programs Biomed*. 2020;**185**:105167. [PubMed ID: 31715333]. https://doi.org/10.1016/j.cmpb.2019.105167.

99. Aiassa S, Ny Hanitra I, Sandri G, Totu T, Grassi F, Criscuolo F, et al. Continuous monitoring of propofol in human serum with fouling compensation by support vector classifier. *Biosens Bioelectron*. 2021;**171**:112666. [PubMed ID: 33069958]. https://doi.org/10.1016/j.bios.2020.112666.

100. Pasma W, Wesselink EM, van Buuren S, de Graaff JC, van Klei WA. Artifacts annotations in anesthesia blood pressure data by man and machine. *J Clin Monit Comput*. 2021;**35**(2):259-67. [PubMed ID: 32783094]. [PubMed Central ID: PMC7943498]. https://doi.org/10.1007/s10877-020-00574-z.

101. Hever G, Cohen L, O'Connor MF, Matot I, Lerner B, Bitan Y. Machine learning applied to multi-sensor information to reduce false alarm rate in the ICU. *J Clin Monit Comput*. 2020;**34**(2):339-52. [PubMed ID: 30955160]. https://doi.org/10.1007/s10877-019-00307-x.

102. Unal A, Arsava EM, Caglar G, Topcuoglu MA. Alarms in a neurocritical care unit: a prospective study. *J Clin Monit Comput*. 2022;**36**(4):995-1001. [PubMed ID: 34043136]. [PubMed Central ID: PMC8156574]. https://doi.org/10.1007/s10877-021-00724-x.

103. Liu Y, Liu C, Gao M, Wang Y, Bai Y, Xu R, et al. Evaluation of a wearable wireless device with artificial intelligence, iThermonitor WT705, for continuous temperature monitoring for patients in surgical wards: a prospective comparative study. *BMJ Open*. 2020;**10**(11). e039474. [PubMed ID: 33208327]. [PubMed Central ID: PMC7677341]. https://doi.org/10.1136/bmjopen-2020-039474.

104. Convertino VA, Wampler MR, Johnson M, Alarhayem A, Le TD, Nicholson S, et al. Validating clinical threshold values for a dashboard view of the compensatory reserve measurement for hemorrhage detection. *J Trauma Acute Care Surg*. 2020;**89**(2S Suppl 2):S169-74. [PubMed ID: 31972755]. https://doi.org/10.1097/TA.0000000000002586.

Table 2. Category A: Neurocritical Care

No.	Study	Goal	Type of Brain Injury	Type of Anesthesia	Induction Drug(s)	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance	Interpretable?
1	Xin et al. (2021) (9)	To evaluate the value of propofol anesthesia for brain protection of patients undergoing craniotomy evacuation of the hematoma	Cerebral hemorrhage	General	Sufentanil and cisatracurium	Request	100/Yantai Yuhuangding Hospital	Extracted features from diffusion tensor imaging images that are a special form of MRI by residual block	DL super-resolution	Multiscale residual network (for experimental group)	FA values, NHISS scores, brain metabolism indexes at some time points, NSE and S100β protein levels, and the probability of postoperative complications in the corticospinal tract of the hind limb of the internal capsule of the affected side were better in the experimental group than in control group	No
2	Farzaneh et al. (2021) (12)	To predict long-term functional outcomes of TBI patients using available data	Traumatic brain injury (TBI) or "Silent Epidemic"	-	-	Request	881 (M: 65, F: 224)/ProTECT III dataset	18 EHR variables and medical history from which three sets of features are extracted, including all candidate variables, excluding non-robust variables, and excluding non-robust and counterintuitive variables	XGBoost	XGBoost on "All candidate variables" feature set	AUC: 0.809, accuracy: 75.3%, F1-score: 70.5%, sensitivity: 70.1%, specificity: 79.1%, precision: 70.9%	Yes, using Shapely
3	Koch et al. (2021) (15)	To ascertain potential insights into pathological mechanisms of injury after aSAH	Aneurysmal subarachnoid hemorrhage	-	-	Unavailable	81 (M: 32, F:49) cerebro spinal fluid samples/not reported	Patient demographic and clinical characteristics, including World Federation of Neurological Surgeons grade, modified Fischer score, means of treatment, and need for permanent CSF diversion	Elastic Net ML and orthogonal partial least squares-discriminant analysis	EN and OPLS-DA	EN ML and OPLS-DA analysis identified 8 and 10 metabolites, respectively	No
4	Schweingruber et al. (2022) (4)	To predict critical phases of intracranial hypertension in patients with invasive ICP measurement	Evolution of ICP	-	-	External datasets are available at PhysioNet.org, and local datasets are available upon request.	3978 including local dataset: ICP-ICU dataset (1346) and external datasets: (MIMIC-III (998) and eICU (1634))/not reported	Descriptives (age, weight, height, diagnosis) and most common and frequent features in all databases (vital signs, laboratory, medication, blood-gas analysis)	LSTM	LSTM	Using LSTM in this study had good results.	No
5	Bernabei et al. (2021) (13)	To present a real-time alerting and monitoring system for epilepsy and seizures that dramatically reduces the amount of manual electroencephalogram review	Epilepsy and seizures	-	Thiopental, midazolam, ketamine	Available	97 (M: 44, F:53)/ICUs at the University of Pennsylvania Health System	Continuous EEG signals: Power in the delta, theta, alpha, beta frequency bands, signal line length, wavelet entropy, statistical features of the signal, the mean value of the upper signal envelope of the electroencephalogram waveform	RF	RF	Mean seizure sensitivity: 84% (cross-validation) and 85% (testing), mean specificity: 83% (cross-validation) and 86% (testing)	Yes, using RF.
6	Narula et al. (2021) (10)	To detect bursts in EEG and generate burst-per-minute estimates for the purpose of monitoring the sedation level in an ICU	Intracranial hemorrhage	-	Isoflurane	Unavailable	29 (M: 16, F:13)/Neurocritical Care Unit, University Hospital Zurich	Continuous EEG signals: Distance between covariance matrices	BSUPP (new unsupervised burst suppression detection algorithm)	BSUPP	Mean absolute error in bursts per minute: 0.93, average of Sensitivity: 81%, average of specificity: 81%, AUROC: 0.82, average NPV: 97%	No
7	Fumeaux et al. (2020) (14)	To create a seizure-detection approach	Spontaneous seizures	-	-	Unavailable	112/focal epilepsy dataset and multifocal epilepsy dataset	(Continuous EEG) cEEG signals: RMS of signal, coastline, skewness, kurtosis, autocorrelation function, Hjorth parameters (activity, mobility, complexity of EEG signal), maximal cross-correlation, and extra	GLM	GLM	AUROC: 0.890 latency to detection: Under 5 seconds for over 80% of seizures and under 12 seconds for over 99% of seizures	Yes, using the logit link function

8	Farzaneh et al. (2020) (11)	To segment and assess the severity of subdural hematoma for patients with TBI	TBI	Sedation	Sedation with Propofol or dexmedetomidine, analgesia with fentanyl	Unavailable	11/Michigan Medicine Neurological Intensive Care Unit or Emergency Department	Computed tomography scans: Age, location-based (radial distance, Azimuth angle, elevation angle, distance to skull), histogram-based (minimum, maximum, average, SD, skewness, kurtosis, entropy), filtering-based (Gabor, Laplacian of Gaussian), deep features	RF	RF+Post-processing	Recall: 98.81%, specificity: 92.31%, F1-score: 98.22%	No
9	Elmer et al. (2020) (16)	To detect early post-cardiac-arrest brain injury phenotypes	Hypoxic-ischemic brain injury	Sedation	Sedation with propofol or dexmedetomidine, analgesia with fentanyl	Available	1086 (M: 613, F:437)/not reported	Neurological examination, EEG, and brain CT imaging	K-prototypes	K-prototypes	Survival to hospital discharge: 27%	Yes, using the center of clusters

Abbreviations: MRI, magnetic resonance imaging; DL, deep learning; FA, fractional anisotropic; NHISS, National Institute of Health Stroke Scale; NSE, neuron-specific enolase; TBI, traumatic brain injury; EHR, electronic health records; XGBoost, extreme gradient boosting; AUC, area under the curve; aSAH, aneurysmal subarachnoid hemorrhage; CSF, cerebrospinal fluid; ML, machine learning; EN, elastic Net; OPLS-DA, orthogonal partial least squares-discriminant analysis; ICP, intracranial pressure; ICU, intensive care unit; LSTM, long short-term memory; EEG, electroencephalogram; RF, random forest; BSUPP, unsupervised burst suppression detection algorithm; AUROC, area under the receiver operating characteristics; NPV, negative predictive value; GLM, generalized linear model; CT, computed tomography.

Table 3. Category B: Pain Management

No.	Study	Goal	Type of Surgery	Type of Anesthesia	Induction Drug(s)	Evaluation Pain Index	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance
1	Tan et al. (2021) (17)	Identifying parturients at increased risk of breakthrough pain during labor epidural analgesia	Parturition	Regional	Fentanyl and ropivacaine.	-	Unavailable	20798/KK Women's and Children's Hospital, a tertiary obstetric hospital	Maternal age, race/ethnicity, BMI, ASA PS score, parity, twins, pre-neuraxial analgesia pain score, pre-neuraxial analgesia cervical dilation, post-neuraxial analgesia highest pain score (0-10), analgesia use prior to neuraxial analgesia, neuraxial technique, combined spinal-epidural, number of neuraxial attempts, median, neuraxial procedure total time, median, depth to epidural space, median, length of catheter in epidural space, median and etc.	RF, XGBoost & LR	LR	Sensitivity: 69.4%, specificity: 73.3%, PPV: 30.1%, NPV: 93.5%
2	Barry et al. (2021) (21)	To investigate the incidence and factors associated with rebound pain in patients who received a PNB for ambulatory surgery	Ambulatory surgeries	Local (peripheral nerve block)	Ropivacaine (or bupivacaine) with lidocaine.	Numerical rating scale (NRS)	Unavailable	972/Hospital databases Draagerwerk AG & Co	Age, BMI, gender, surgery duration, local anesthetic volume, local anesthetic dose, sensory block duration, motor block duration, ASA physical status, surgical site, surgical site (specific), surgery type, general anesthesia, peripheral nerve block type, local anesthetic drugs, analgesia adjuncts, postoperative NSAID use, postoperative acetaminophen use, postoperative opioid use	Univariate linear regression, multivariable LR, logistic model tree attribute-selected classifier	Logistic model tree attribute-selected classifier	ROC: 0.6
3	Choi et al. (2021) (20)	Develop a new analgesic index to objectively assess pain in conscious patients.	Breast, colorectal, hepatobiliary, stomach, thyroid	General	Propofol and remifentanil	Spectrogram-CNN index	Unavailable	100 (M:44, F: 56)/not reported	Photoplethysmogram spectrograms, gender, age, height, weight, ASA PS_a, type of surgery, Postoperative pain intensity at $PACU_b$	CNN	CNN	AUC: 0.76 balanced accuracy: 71.4%, sensitivity: 68.3%, specificity: 73.8%
4	Gonzalez-Cava et al. (2020) (18)	Evaluate the suitability of the analgesia Nociception index as a guidance variable to replicate the decisions made by the experts when a modification of the opioid infusion rate is required.	Cholecystectomy surgery	General	Remifentani and propofol	Analgesia Nociception index (ANI)	Unavailable	17 (M: 4, F:13)/Hospital Universitario de Canarias	Feature vector proposal 1: Hemodynamic information (SP, SP5, SP10 DP, DP5 DP10 HR, HR5, HR10 Remi, Remi5, Remi10) Feature vector proposal 2: Minimum ANI information (SP, SP5, SP10, DP, DP5 DP10, HR, HR5, HR10, Remi, Remi5, Remi10, and extra,	KNN, DT, LDA, SVM, LR, ensemble classifiers	SVM	Accuracy: 86.21%, precision: 86.11%, recall: 91.18%, specificity: 79.17%, AUC: 0.89 Kappa index: 0.71
5	Teichmann et al. (2020) (22)	Detection of dental pain sensation based on cardiorespiratory signals using a machine learning classifier	Dental treatment	General	-	-	Unavailable	20 (M: 16, F:4)/Department of Prosthodontics and Biomaterials-Center of Implantology, Medical Faculty, RWTH Aachen University	Frequency spectral bins, levels of the discrete wavelet transform, average height, maximum deviation in height, average pulse beat-to-beat time, maximum deviation in beat-to-beat times, average area, the maximum deviation of areas, the average ratio between pulse width and height, the maximum deviation of the ratio between pulse width and height	RF	RF	Sensitivity: 87%, specificity: 63%, AUC: 0.828
6	Meijer et al. (2020) (19)	To reduce postoperative pain using Nociception level-guided opioid dosing during general anesthesia	Abdominal surgery	General	Fentanyl & sevoflurane	Nociception level (NOL) index	Unavailable	50 (M: 22, F:28)/Leiden University Medical Centre, Alrijne Hospital	Age, gender, weight, height, BMI, MAP, HR, ASA physical status, general surgery, gynecology, urology	NOL-guided dosing, standard care dosing	NOL-guided Dosing	Median postoperative pain score: 3.2 postoperative morphine consumption (SD): 0.06 (0.07)

Abbreviations: BMI, body mass index; ASA, American Society of Anesthesiology; CSF, cerebrospinal fluid; CSE, combined spinal-epidural analgesia; RF, random forest; XGBoost, extreme gradient boosting; LR, logistic regression; PPV, positive predictive value; NPV, negative predictive value; HER-2, human epidermal growth factor receptor-2; DT, decision tree; GB, gradient boosting; LightGBM, light gradient boosting machine; AUC, area under the curve; PNB, peripheral nerve block; ROC, receiver operating characteristics; CNN, convolutional neural network; ASA PS, American Society of Anesthesiologists Physical Status; PACU, post anesthesia care unit; ANI, analgesia nociception index; KNN, K-nearest neighbor; LDA, linear discriminant analysis; SVM, support vector machine; NOL, nociception level; MAP, mean arterial pressure; HR, heart rate.

Table 5. Subcategory D_1: Perioperative

No.	Subcategory	Study	Goal	Type of Surgery	Type of Anesthesia	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance	Interpretable?
1	D1	Maheshwari et al. (2020) (27)	To evaluate the performance of the hypotension (MAP < 65 mmHg for at least 1 min) prediction index algorithm derived from non-invasive arterial pressure waveforms in moderate-to-high-risk non-cardiac surgical patients	Non-cardiac	General	Unavailable	305/ClearSight, Edwards Lifesciences	Waveform features and the patient demographics, including age, gender, height, and weight.	HPI algorithm for the 5, 10, and 15 minutes of prediction time points before each hypotensive episode for blinded arm, unblinded arm, combined groups	HPI algorithm for 5 min prediction time point before hypotension episode for blinded arm	AUC: 0.94, sensitivity: 86%, specificity: 87%	No
2	D1	Li et al. (2021) (28)	Prediction of post-induction hypotension (SBP < 90 mmHg or MBP < 65 mmHg) in patients undergoing cardiac surgery	Cardiac	General	Unavailable	3030/The Second Affiliated Hospital of Hainan Medical University	Preoperative variables including age, gender, BMI, underlying disease, EuroSCORE I, and ASA score; experimental findings including hemoglobin, serum creatinine, and total bilirubin; data on the patient's preoperative medications, such as the use of beta-blockers, insulin, aspirin, intraoperative medications and data on perioperative blood pressure	RF	RF	AUC: 0.843	Yes, using the interpretability of RF
3	D1	Frassanito et al. (2020) (29)	To assess the diagnostic ability of Hypotension Prediction Index (HPI) working with non-invasive ClearSight system in predicting impending hypotension (MAP < 65 mmHg for > 1 min) in patients undergoing major gynecologic oncologic surgery	Gynecologic oncologic	General	Unavailable	28/Edwards Lifesciences HemoSphere platform	Extracted features from non-invasive arterial pressure waveform of ClearSight	HPI algorithm for the 5, 10, and 15 minutes of prediction time points before each hypotensive episode.	HPI algorithm for 15 min prediction time point before hypotension episode	AUC [95% CI]: 0.95, sensitivity [95% CI]: 85%, specificity [95% CI]: 85%, positive predictive value [95% CI]: 75%, negative predictive value [95% CI]: 91%	No
4	D1	Gratz et al. (2020) (30)	To predict the likelihood of a given patient developing significant hypotension (SBP < 90 mmHg) under spinal anesthesia when undergoing a cesarean section (C/S)	Cesarean	Local	Unavailable	45/not reported	Extracted features from signals using neural network model physiological data, including systole, diastole, mean arterial pressure (MAP), heart rate, and the AS parameter, on a beat-by-beat basis.	NN	NN	AUC: 0.87	No
5	D1	Lee et al. (2020) (31)	To predict hypotension (SBP < 90 mmHg or MBP < 65 mmHg) after tracheal intubation after intubation one minute in advance	Underwent laparoscopic cholecystectomy	General	Unavailable	282/Soonchunhyang University Bucheon Hospital	Totally we had two kinds of features in this study: Raw features and statistical features, including electronic health records (demographic data, comorbidities, baseline) and vital recorder (mechanical ventilation data, bispectral index, anesthetic drug, vasoactive drug administration, Some information about hypotension)	Meta-learning models, such as RF, XGboost, DL models, especially CNN and DNN	Raw features: CNN Statistical features: RF	Accuracy of CNN for raw features: 72.6%, accuracy of RF for statistical features: 74.8%	Yes, using the feature importance of RF
6	D1	Kang et al. (2021) (32)	To predict hypotension (SBP < 90 mmHg or MBP < 65 mmHg) in late Post Induction	Laparoscopic cholecystectomy	General	Available	222/Soonchunhyang University Bucheon Hospital	In this study, 4 feature sets were created by different methods of feature selection, including feature set A (Min				

								(hypotension frequency, Max plasma concentration of propofol), all features (Min effect-site concentration of propofol, Max target concentration of propofol)	Four ML models, including NB, LR, RF, ANN	RF	Accuracy (Feature set C): 79.4%, precision (Feature set C): 81.1%, recall (Feature set B): 84.5%, AUC (Feature set C): 0.842, 95% CI (Feature set C): 0.736-0.948	Yes, using the feature importance of RF
7	D1	Wijnberge et al. (2020) (33)	To predict hypotension (MAP < 65 mmHg for at least 1 min) shortly before it occurs has been developed and validated	Elective noncardiac	General	Unavailable	60 (M: 36, F: 24)/Amsterdam University Medical Centers, Location	Extracted features from the signal (patients based on characteristics divided into intervention and control groups)	HPI algorithm for the 5, 10, and 15 minutes of prediction time points before each hypotensive episode.	HPI algorithm for 15 min prediction time point before hypotension episode	The median time-weighted average of hypotension: 0.10 mm Hg (intervention group); 0.44 mm Hg (control group)	No
8	D1	Solomon et al. (2021) (25)	To predict the occurrence of clinically significant intraoperative bradycardia at time points during an operative course by utilizing available preoperative electronic medical records and intraoperative anesthesia information management system data	Non-cardiac	General	Unavailable	62182/ University of Washington Medical Center	Extracted features from time series signal	Build three models named TP1, TP2 & TP3 by using: GBM & LR	GBM	AUC: 0.89, specificity: 95%, sensitivity: 53%, PPV: 15%, NPV: 99%	Yes, using predictor variables of GBM
9	D1	Jalali et al. (2021) (26)	To predict blood product transfusion requirements for individual pediatric patients undergoing craniofacial surgery	Craniofacial surgery	-	Request	2143/Pediatric Craniofacial Surgery Perioperative Registry	Demographic and preoperative features	Six ML classification and regression models, including RF, AdaBoost, NN, GBM, SVM, Elastic Net methods	GBM	In classification: Sensitivity: 92% ± 3%, specificity: 89% ± 4%, F1-score: 91% ± 4%, AUROC: 0.87 ± 0.03, in regression: MSE: 1.15 ± 0.12, R-squared: 0.73 ±0.02, RMSE: 1.05 ± 0.06	Yes, using the feature ranking of GBM
10	D1	Kim et al. (2021) (34)	To develop and validate practical predictive models for difficult laryngoscopy	-	-	Unavailable	616/Hallym University Chuncheon Sacred Heart Hospital	Age, Mallampati grade, BMI, Sternomental distance, neck circumference	MLP, LR, SVM, RF, XGBoost, LightGBM	LGBM	AUROC: 0.71 Sensitivity: 85%	No
11	D1	Kim et al. (2021) (35)	To predict difficult laryngoscopy of neck circumference and thyromental height	-	General	Request	1677 (M: 925, F: 752)/Hallym University Chuncheon Sacred Heart Hospital	Age, gender, height, weight, BMI, neck circumference, thyromental height	MLP, LR, SVM, RF, XGBoost, LightGBM	RF	AUROC: 0.79 AUPRC: 0.32	No
12	D1	Bollepalli et al. (2021) (37)	To improve life-threatening arrhythmia detection in the ICUs	-	-	Request+https://physionet.org/content/challenge-2015/1.0.0/	410/ICUs of Massachusetts General Hospital and PhysioNet	Deep features+ECG, blood pressure, PPG features (periodicity measure, sharpness measure, correlation measure, peak height stability measure, and extra.	Hybrid CNN	Hybrid CNN	Accuracy: 87.5% ± 0.5%, score: 81% ± 0.9%, evaluation on PhysioNet 2015 Challenge database: Accuracy: 84.3%, score: 93.9%	No
13	D1	Yeh et al. (2021) (38)	To classify ECG image types to assist in anesthesia assessment	-	-	Available	54190/MIT-BIH Arrhythmia Database	2D ECG images	ResNet, AlexNet, SqueezeNet	ResNet	Accuracy: 97%, recall: 97%, precision: 97, F1-score: 97%, Kappa statistics: 0.96	No
14	D1	Hadjipavlou et al. (2021) (39)	Exploring elements of synaptic transmission, looking for possible contributions to the									

			anesthetized EEG	-	General	Unavailable	Not reported/ Oxford University Clinical Academic Graduate School	Simulated electrocorticography: Alpha band at rest, loss of frequencies at induction, alpha and slow wave bands at maintenance, and broad spectral activity at emergence. AG, anesthetic GABA	Hodgkin-Huxley-type NN computer simulation	Hodgkin-Huxley-type NN computer simulation	-	No
15	D1	Mathis et al. (2020) (36)	Identifying patients ultimately diagnosed with heart failure with reduced ejection fraction following surgery using preoperative and intraoperative data	Noncardiac surgery	General	Unavailable	67697 (M: 32200, F: 35497)/ Multicenter Perioperative Outcomes Group (MPOG) database+Epic Systems	628 preoperative and 1195 intraoperative features	L1 Regularized LR, RF, XGBoost	XGBoost	AUROC: 0.873, AUPRC: 0.040, accuracy: 80.82%, sensitivity: 80.84%, specificity: 80.82%, PPV: 1.78%, NPV: 99.90%	Yes, using the feature importance

Abbreviations: MAP, mean arterial pressure; HPI, hypotension prediction index; AUC, area under the curve; SBP, systolic blood pressure; MBP, mean blood pressure; BMI, body mass index; ASA, American Society of Anesthesiology; RF, random forest; AS, arterial stiffness; NN, neural network; XGboost, extreme gradient boosting; DL, deep learning; CNN, convolutional neural network; DNN, deep neural network; PIH, post induction hypotension; ML, machine learning; NB, naïve bayes; LR, logistic regression; GBM, gradient boosting machine; PPV, positive predictive value; NPV, negative predictive value; SVM, support vector machine; AUROC, area under the receiver operating characteristics; MSE, mean square error; RMSE, root mean square error; MLP, multi-layer perceptron; LGBM, light GBM; AUPRC, area under the precision-recall curve; ICU, intensive care unit; ECG, electrocardiogram; PPG, photoplethysmography; HR, heart rate; EEG, electroencephalogram; GABA, gamma-aminobutyric acid.

Table 6. Subcategory D_2: Postoperative

No.	SubCat.	Study	Goal	Type of Surgery	Type of Anesthesia	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance	Interpretable?
1	D2	Racine et al. (2021) (42)	To predict delirium in a rigorous and well-characterized, prospective, observational cohort study of delirium	Elective non-cardiac including	-	Unavailable	560/Beth Israel Deaconess Medical Center, Brigham and Women's Hospital, and Hebrew SeniorLife	Medical records includig: Surgical procedure, anesthesia type and duration, baseline diagnoses and comorbidity, abnormal laboratory results, development of delirium, precipitating factors for delirium (e.g., medications, iatrogenic events, catheters, or physical restraints), postoperative complications, and intercurrent illnesses	GB, Cross-validated LR, NN, RF, and Regularized Regression (least absolute shrinkage and selection and ridge regularization) & two ensemble approaches	Cross-validated LR for full feature set	AUC: 0.7; Sensitivity: 46%; Specificity: 81%; PPV: 43%; NPV: 83%	No
2	D2	Lu et al. (2021) (52)	To identify patients requiring admission following elective anterior cruciate ligament reconstruction	Non-elective	Different type of anesthesia were used, including: Epidural, General, MAC/IV sedation, Regional, Spinal, Operative time	Unavailable	4709/The ACS National Surgical Quality Improvement Program database	age, Gender, BMI, functional status, level of dyspnea, ASA Physical Status Classifcation, location from which patient was admitted, anesthesia type, operative time, admission quarter, diabetes mellitus, congestive heart failure, chronic obstructive pulmonary disease, smoking history, preoperative sepsis, preoperative use of a ventilator, ascites, wound infection, weight loss>10%, etc.	RF, XGBoost, LDA, AdaBoost & An additional model was produced as a weighted ensemble of the four fnal algorithms	Ensemble model	AUC: 0.76	Yes
3	D2	Lee et al. (2021) (40)	To learn patterns related to risk of in-hospital mortality for patients undergoing surgery under general anesthesia	-	General	Unavailable	59985/UCLA Medical Center's Perioperative Data Warehouse	Medical information including: Age, Estimated blood loss, Presence of arterial line, Presence of pulmonary artery line, Presence of central line, ASA score and other & Healthcare Cost and Utilization Project (HCUP) Code Descriptions including: UPPER GASTROINTESTINAL ENDOSCOPY, BIOPSY 3864, COLONOSCOPY AND BIOPSY, LAMINECTOMY, EXCISION INTERVERTEBRAL DISC and other	Generalized Additive Models with NN (GAM-NN) & LR	GAM-NN	AUC: 0.921; AP: 17.6%	Yes, using Interpretable model (GAM-NN)
4	D2	Schenk et al. (2021) (44)	To investigate the effect of Hypotension Prediction Index-guided intraoperative haemodynamic care on depth and duration of postoperative hypotension	Elective noncardiac	General	Unavailable	54/Amsterdam University Medical Centers	Extracted features from the invasive Blood Pressure signals	HPI algorithm	HPI algorithm	Intraoperative HPI-guided haemodynamic care did not reduce the TWA of postoperative hypotension	No
5	D2	Tan et al. (2021) (45)	Prediction of early phase postoperative hypertension requiring the administration of intravenous vasodilators after carotid endarterectomy	-	General	Unavailable	367/Huashan Hospital of Fudan University	Patient demographics, CEA procedure details, parameters of laboratory examination, imaging study & perioperative blood pressure	GBR Trees	GBR Trees	Average AUC: 0.77; Average Specificity: 52%; Sensitivity ~ 90%	Yes, using feature importance of GBRT
6	D2	Lu et al. (2021) (53)	To predict cost after anterior cruciate ligament reconstruction	Ambulatory ACLR	Different types of anesthesia were used, including: MAC/IV sedation, Local anesthesia, General anesthesia & Regional anesthesia	Unavailable	7311/New York State Ambulatory Surgery and Services database	Features included in initial models consisted of patient characteristics (age, Gender, insurance status, income, medical comorbidities as classified by the Clinical Classifications Software diagnosis code) as well as intraoperative variables (type of anesthesia and procedure-specific factors)	Four ML models including: RF, XGBoost, Elastic Net Penalized Regression & SVMs with radial kernels	RF	Accuracy: 87.8%; AUC: 0.848; Calibration and the Brier score: 20.8%	Yes, using interpretability of RF
7	D2	Palla et al. (2022) (43)	To predict hypotension in the recovery area better than clinicians using readily available clinical information	Different type of surgery like Orthopaedic, General, Urology, ENT, etc	-	Unavailable	121904/Two UW hospitals	Demographics data, Procedure details, Comorbidities, Vitals, Drugs & other	GBRT	GBRT	AUROC: 0.82; AUPRC: 0.4	Yes, using ShAP Value
8	D2	Jeong et al. (2021) (46)	To predict postoperative complications, major adverse cardiac events, for patients who underwent any type of surgery	Any type of surgery	General	Request	586/Soonchunhyang university Seoul hospital	pre-op EMR features: demographic values (e.g., height, weight, Gender, age, BMI), several pre-op evaluation results (e.g., EF, PFT), pre/post hemodialysis evaluations (e.g., Na, K, Cl), and comorbidities (e.g., hypertension, atrial fibrillation) peri-op features: Anesthesia-related values (e.g., ASA, EM emergency operation,				

								anesthesia method), and other operation-related values (e.g., anesthesia time, operation time, infusion of crystalloid or colloid) text features: Generated by applying NLP techniques to preanesthetic assessment documents	SVM, DT, RF, Gaussian NB, ANN, LR, XGBoost	RF	F1-score: 79.7%	Yes, using Recursive Feature Elimination (RFE) and K-best
9	D2	Qian et al. (2021) (50)	To assess the significance of operative timing on classifying surgical complications	Different type of surgery like Obstetric, Gynecological, Liver, etc.	All types of anesthesia	Request	107481(M:55515,F:51966)/University-affiliated, tertiary teaching hospital	Date and Time the Surgery, Duration of Surgery, Length of Stay, Surgical Discipline, Patient Age and Gender, Admission and Discharge Consultation Summaries, Preoperative Comorbidity (if any), Postoperative Complications (if any)	LR, NB CART, RF, GBDT, AdaBoost, XGBoost, LightGBM, CatBoost	XGBoost	Accuracy: 95%; Precision: 96%; Recall: 94%; F1-score: 95%; AUC: 0.98	Yes, using interpretable classifiers
10	D2	Chelazzi et al. (2021) (41)	To identify patients at risk for postoperative complications	Different type of surgery like Breast surgery, Dental surgery, Endocrine surgery, etc.	-	Request	560/Tertiary care teaching hospital of Careggi (Azienda Ospedaliero-Universitaria di Careggi)	Patients comorbidity factors: Abnormal ECG (lef bundle branch block, lef ventricular hypertrophy, repolarization abnormalities, non-sinus rhythm), Untreated hypertension or hypertension not controlled by medical therapy, Previous thromboembolism, Stable or controlled angina, Previous myocardial infarction with no clinical or diagnostic evidence of residual ischemia, Compensated heart failure or previous heart failure, Diabetes mellitus, and etc.	Single Layer Feedforward Network with the training algorithm.	DEC	Average Classifcation Accuracy: 90%; Balanced Accuracy: 90.45%; Sensitivity: 88.9%; Specificity: 90.2%; PPV: 61.5%; NPV: 97.9%	No
12	D2	Bishara et al. (2022) (5)	To develop a postoperative delirium risk prediction model	Different type of surgery like Neurological Surgery, Orthopedics Surgery, General Surgery, etc.	-	Request.	24885(M:12276,F:12609)/Moftt-Long Hospital, Mission Bay Hospital	Demographics, Comorbidities, Nursing Assessments, Surgery Type, and other preoperative pre-operative electronic health data	NN, XGBoost, Clinician-Guided Regression, ML Hybrid Regression, AWOL-S	XGBoost	AUC-ROC: 0.851	Yes, using XGBoost
13	D2	Bai et al. (2020) (47)	To provide clinical data for the prevention of postoperative cerebral infarction and myocardial infarction	-	General	Request	443(M:351,F:92)/Peking University Third Hospital	Demographic Data, Previous Medical History, Degree of Neck Vascular Stenosis, Blood Pressure at time points during the perioperative period, the Time of Occlusion, whether to Place the Shunt, and the time of Hospital Stay, whether to have Cerebral Infarction and Myocardial Infarction	SVM, DT, RF, ANN, Quadratic Discriminant Analysis, XGBoost	XGBoost	Accuracy: 94%	No
14	D2	Ko et al. (2020) (48)	Identification of preoperative risk factors for postoperative acute kidney injury	Knee arthroplasty	General, Spinal	Unavailable	5757(M:682,F:5075)/not reported	Preoperative serum creatinine levels, use of TXA, general anesthesia, use of RAASis, ASA class, and Gender	GBM	GBM	AUC: 0.78	No
15	D2	Suhre et al. (2020) (51)	Association of cannabis use with a small increase in the risk of postoperative nausea and vomiting	-	General	Available	43633/University of Washington Medical Center, Harborview Medical Center	Age, ASA, Outpatient, Gender, Non-smoker, Prior PONV/Motion Sickness, Procedure Duration, Exposed to Nitrous Oxide, Surgery Higher Risk for Nausea, Total Number of Prophylactic Agents, PACU Opioids, Apfel Score	Bayesian Additive Regression Trees	Bayesian Additive Regression Trees	Mean Relative Risk: 1.19	No
16	D2	Cao et al. (2020) (49)	To explore whether serious postoperative complications of bariatric surgery recorded in a national quality registry can be predicted preoperatively	Bariatric Surgery	General	Unavailable	44061/Scandinavian Obesity Surgery Registry	5 continuous features (age, hemoglobin A1c, BMI, WC, and operation year) and 11 dichotomous features (Gender sleep apnea hypertension diabetes dyslipidemia dyspepsia depression musculoskeletal pain previous venous thromboembolism revisional surgery and the outcome, serious postoperative complications)	MLP, CNN, RNN	CNN	AUC: 0.57	No

Abbreviations: GB, gradient boosting; LR, logistic regression; NN, neural network; RF, random forest; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; BMI, body mass index; ASA, American Society of Anesthesiology; XGBoost, extreme gradient boosting; LDA, linear discriminant analysis; GAM, generalized additive model; AP, average precision; HPI, hypotension prediction Index; TWA, time-weighted average; CEA, carotid endarterectomy; GBR, gradient boosted regression; GBRT, gradient boosted regression trees; ACLR, anterior cruciate ligament reconstruction; ML, machine learning; SVM, support vector machine; AUROC, area under the receiver operating characteristics; AUPRC, area under the precision-recall curve; ICU, intensive care unit; EMR, electronic medical record; NLP, natural language processing; DT, decision tree; NB, naïve bayes; ANN, artificial neural network; GBDT, gradient boosted decision trees; ECG, electrocardiogram; ROC, receiver operating characteristics; TXA, tranexamic acid; RAASis, renin–angiotensin–aldosterone system inhibitors; GBM, gradient boosting machine; PONV, postoperative nausea and vomiting; PACU, post anesthesia care unit; WC, waist circumference; MLP, multi-layer perceptron; CNN, convolutional neural network; RNN, recurrent neural network.

Table 7. Subcategory D_3: Critical Care

No.	SubCat.	Study	Goal	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance	Interpretable?
1	D3	Magunia et al. (2021) (54)	To stratify patient risk and predict ICU survival and outcomes	Request	1039(M:853,F:333)/27 German hospitals	A total of 49 variables were used for the ML models, including: Demographic data, Past medical history, Previous medications, Current illness data, Laboratory values as well as outcome data	Explainable Boosting Machine (EBM), EBM with 10 interactions, SVC & RF	EBM with 10 interactions	Balanced Accuracy: 64%, PR-AUC: 0.81	Yes, using interpretable model
2	D3	Hu et al. (2021) (59)	To incorporate key variables into a parsimonious model for electroencephalographic seizure prediction in critically ill children	Unavailable	719/Research Electronic Data Capture database	Clinical data included age, Gender, prior neurodevelopmental disorders, medications, CEEG indication, hospital and PICU admission and discharge dates, presence of clinically evident seizures prior to CEEG, acute encephalopathy category (epilepsy-related, acute structural, or acute non-structural) based on the primary presenting problems/diagnoses available at the time of admission, and mental status (comatose or not baseline or not)	RF, Least Absolute Shrinkage and Selection Operator & DL Important Features	RF	Training Accuracy: 96.3%; Validation Accuracy: 74%; AUROC: 0.706; F1-score: 73.2%	Yes, using ranking algorithm based on the relative importance
3	D3	Cherifa et al. (2021) (56)	To predict simultaneously the Mean Arterial Pressure and the Heart Rate	Available	22247(M:1424,F:884)/ MIMICIII waveform matched subset from the five ICUs of Boston's Beth Israel deaconess medical center	Patients characteristics (age, gender, ...), Initial severity scores (SOFA, SAPS-II), Type of intensive care unit, Treatment (sedation, vasopressors, mechanical ventilation) & Physiologic signals (pulse, oximetry, heart rate, systolic arterial pressure, mean arterial pressure and diastolic arterial pressure)	Multi-task Learning Physiological Deep Learner (MTL-PDL) & Single-task Learning Physiological Deep Learner (STL-PDL)	MTL-PDL	RMSE of MTL-PDL was less than RMSE of STL-PDL	Yes
4	D3	Moghadam et al (2020) (57)	To predicts hypotension up to 30 min in advance based on the data from only 5 min of patient physiological history in ICU	Unavailable	1000(M:604,F:396)/MIMIC III database	A set of 33 scalar features are used to represent each data point. At each data point, including: Arterial blood pressure, Heart rate, Systolic blood pressure, Diastolic blood pressure, Respiration rate, Peripheral capillary oxygen saturation, Pulse pressure, Mean arterial pressure, Cardiac output, MAP to HR ratio, and etc.	LR, a variety of SVM algorithms, and KNN with different kernels	LR	Accuracy: 94%, sensitivity: 85%, specificity: 96%, PPV: 81%	Yes, using feature importance
5	D3	Cherifa et al. (2021) (58)	To predict an Acute hypotensive episodes, 10 minutes in advance	Available	1320/MIMIC II database(1151) & External dataset from Lariboisière hospital was used for external validation(169)	Age, Gender, type of care unit, severity scores, and time-evolving characteristics such as Mechanical ventilation, vasopressors, or sedation medication as well as features extracted from physiological signals: heart rate, pulse oximetry, and arterial blood pressure	For Random partial sample: Bayesian Generalized Linear Regression, XGBoost, Gradient Boosting, Interaction LR, LR, NN, Penalized LR, RF, Recursive Partitioning, Discrete Super learner and Super Learner & For full sample: Generalized Linear Mixed algorithms via PQL, Generalized Linear Mixed algorithms via ML, Linear regression using Generalized Least Squares, Discrete Super learner and Super Learner	For the first task, that is, AHE prediction based on 1 random period per patient (random partial sample): RF & For AHE prediction based on all periods (full sample): The Generalized Linear Mixed ensemble weight of 0.70	RF: BS: 0.086 & The Generalized Linear Mixed ensemble: 0.082	No
6	D3	Yun et al. (2021) (55)	To predict in-hospital death of critically ill patients with considerable accuracy and identify factors contributing to the prediction power	Request	1384/Surgical Intensive Care Unit of their institution	Demographic variables (Age, gender, BMI, ...), Disease-specific variables (Disease diagnosis, origin, ...), Surgical variables (Type of surgery, operation name, ...), Laboratory variables (Blood gas analysis, WBC, ...) & Hemodynamic variables (Use of inotropes and use of vasopressors)	DT, NN, NB, RF and Hellinger Distance Estimates	RF	F1-score: 84%; Precision: 78%; Recall: 90%; AUC: 0.77	Yes
7	D3	Chang et al. (2022) (60)	To predict ICU admission of patients with Myasthenia Gravis	Request	228/Shin-Kong Wu Ho-Su Memorial Hospital	Medical records including information on the age, Gender, age at diagnosis, disease duration, autoantibodies present, medications used, maximum dosage of corticosteroid before admission, thymic histology, history of thymectomy, treatment during hospitalization and length of ICU admission	Classification and regression tree, C4.5 & C5.0	C5.0 DT	Accuracy Mean (SD): 94.2%; Sensitivity Mean (SD): 99.4%; Specificity Mean (SD): 63.9%; AUC Mean (SD): 0.814; F1-score Mean (SD): 96.7%	Yes
8	D3	Hayasaka et al. (2021) (61)	To classify intubation difficulties from the patient's facial image	Request	202(M:92,F:110)/Yamagata University Hospital	Facial Images	Classification and regression tree, C4.5 & C5.0	CNN	Accuracy: 80.5%; Sensitivity: 81.8%; Specificity: 83.3%; AUC: 0.864	No

9	D3	Wu et al. (2021) (62)	To investigate the association between culture positivity during admission and long-term outcome in critically ill surgical patients	Request	6748/Taichung Veterans General Hospital, Taiwanese National Health Insurance Research Database	Age, Gender, BMI, Comorbidities, Severity Score, Shock, Early Fuid Overload, Receiving Mechanical Ventilation, the Need of Renal Replacement Therapy for Critical Illness	Log-rank test + multivariable Cox proportional hazards regression model	Log-rank test + Multivariable Cox proportional hazards regression model	Hazard Ratio: 1.579	No
10	D3	Ling et al. (2021) (63)	Investigate the relationship between the red cell distribution width and the prognosis of patients with Sepsis-associated thrombocytopenia	Request	809(M:444,F:365)/MIMIC-III database	Age, Gender, Hypertension, Diabetes, Stroke, Heart diseases, Red Cell Distribution Width, Hemoglobin, Hematocrit, White Blood Cells, Platelet count, Prothrombin Time, Activated Partial Thromboplastin Time, Lactate, Sequential Organ Failure Assessment score	XGBoost	XGBoost	Sensitivity: 70%; Specificity: 57%; AUC: 0.646	Yes, using SHapley Additive exPlanations
11	D3	Sarton et al. (2021) (64)	Investigate the association between early brain MRI data and functional outcomes of patients with severe herpes simplex encephalitis at 90 days after ICU admission	Unavailable	138(M:75,F:63)/34 ICUs in France	Patient's history, clinical, laboratory, and brain electrophysiologic data	Multivariable LR	Multivariable LR	AUC: 0.87; Goodness of fit (Hosmer and Lemeshow test): 0.75; Accuracy: 81.4%	No
12	D3	Elmer et al. (2020) (65)	To predict Cerebral Performance Category using longitudinal data after cardiac arrest	Unavailable	1010(M:626,F:384)/not reported	EEG data	Group-Based Trajectory Modeling (GBTM)-unadjusted, GBTM-Ocov, GBTM-Risk, GBTM Ocov+Risk, K-means-unadjusted, K-means-Adjusted, Bayesian regression	GBTM-Risk	Sensitivity: 38.3%	Yes, using Centers of Clusters

Abbreviations: ICU, Intensive Care Unit; ML, machine learning; EBM, explainable boosting machine; SVC, support vector classifier; RF, random forest; PR-AUC, precision recall area under the curve; CEEG, continuous electroencephalogram; PICU, Pediatric Intensive Care Unit; DL, deep learning; AUROC, area under the receiver operating characteristics; MTL-PDL, multi-task learning physiological deep learner; STL-PDL, single-task learning physiological deep learner; RMSE, root mean square error; MAP, mean arterial pressure; HR, heart rate; RR, respiratory rate; ECG, electrocardiogram; ABP, arterial blood pressure; Resp, respiration rate; SpO_2, peripheral oxygen saturation; PP, pulse pressure; CO, cardiac output; LR, logistic regression; SVM, support vector machine; KNN, K-nearest neighbor; PPV, positive predictive value; XGBoost, extreme gradient boosting; NN, neural network; AHE, acute hypotensive episodes; BS, brier score; BMI, body mass index; DT, decision tree; NB, naïve bayes; AUC, area under the curve; CNN, convolutional neural network; MRI, magnetic resonance imaging; EEG, electroencephalogram; GBTM, group-based trajectory modeling.

Table 10. Category G: Depth of Anesthesia

No.	Study	Goal	Type of Anesthesia	Induction Drug(s)	Depth of Anesthesia Levels	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance
1	Abel et al. (2021) (74)	To construct classification models for real-time tracking of anesthesia unconscious state during anesthesia	General	Propofol & Sevoflurane	Consciousness & Unconsciousness	Request	Not reported/ Massachusetts General Hospital	The feature sets used were the multi-tapered EEG spectral power, EEG bendwise power, the first three principal component scores of the multitaper spectrogram, the linear discriminant score of the multitaper spectrogram (LDA, with supervised learning performed by including the labels), and the first ten principal component scores of a set of features generated by a deep CNN.	Use the below algorithms in 3 ways: Without HMM, with 2-State-HMM, and With 6-State-HMM. The algorithms are Multitaper Spectrogram, BWP, PCA, LDA, and CNN.	LDA+HMM2	AUC ~ 0.99
2	Dubost et al. (2021) (89)	To predict and assess states based on four physiological variables: Heart Rate, Mean Blood Pressure, Respiratory Rate, AA Inspiratory Concentration	General	Propofol & Ketamine	Awake, The Loss of Consciousness (LOC), The anesthesia, The Recovery of Consciousness (ROC), Emergence.	Unavailable	30 (M: 20, F: 10)/Begin military teaching hospital	Heart Rate, Systolic arterial blood pressure, Diastolic arterial blood pressure, Mean arterial blood pressure, Saturated percentage of dioxygen, End-tidal carbon dioxide, Anesthesia agent, AA expiratory concentration, AA inspiratory concentration, Total minimum alveolar concentration, Fraction inspired of dioxygen, Mean alveolar concentration, Fraction inspired nitrous oxide, End-tidal nitrous oxide, Respiratory rate, BIS, BIS burst suppression ratio, BIS electromyography & Demographic Features (Age, Gender, etc)	HMM	HMM	Error Prediction: 0.18
3	Afshar et al. (2021) (75)	To get EEG signals and continuously predict the BIS	General with few cases receiving Sedation/ Analgesia and Local anesthesia	Propofol and/or Remifentanil	Deep Anesthesia (DA, BIS: 0-40), General Anesthesia (GA, BIS: 40-60), Light Sedation (S, BIS: 60-80) & Awake (W, BIS: 80-100)	Unavailable	176 (M: 102, F: 74)/Department of Anesthesiology and Pain Medicine, Seoul National University Hospital, College of Medicine	Extracted features from EEG signals by DNN age, height, weight, and anesthesia duration	Combinatorial DL structure involving CNN (inspired by the Inception module), Bidirectional LSTM, an Attention Layer	Combinatorial DL structure	AUC: 0.811 ± 0.527, sensitivity: 77.62%, accuracy: 88.71%
4	Zhan et al. (2021) (90)	To distinguish different anesthesia states, providing a secondary tool for DoA assessment	General	Intravenous Midazolam, Propofol, Sufentanil & Cisatracurium	Anesthesia Induction, Anesthesia Maintenance, Anesthesia Recovery	Request	23/Second Affiliated Hospital of the Army Medical University	Four Heart Rate Variability-derived features in the time and frequency domain were extracted from an electrocardiogram, including HRV high-frequency power, Low-frequency power, High-to-low-frequency power ratio, and Sample entropy and age, Height, Weight, BMI, Duration of surgery, Anesthetic management, Maintenance drugs infusion rate, Additional drugs administrated when approaching the end of surgery.	LR, SVM, DT & DNN	DNN	Precision of anesthesia induction: 58.1%, recall of anesthesia induction: 88.1%, precision of anesthesia maintenance: 96%, recall of anesthesia maintenance: 94.7%, precision of anesthesia recovery: 56.6%, recall of anesthesia recovery: 57.8%, classification accuracy: 90.1%
5	Duclos et al. (2021) (76)	To classify different states of AEC and wPLI measure of FC	-	Propofol	Baseline, Light Sedation, Unconscious, Pre-ROC & Recovery	Request	9/not reported	Extracted features from Functional Connectivity time-series signals by AEC & wPLI	Use the below algorithms in 3 ways: With AEC Features, with wPLI Features, and with both AEC and wPLI Features. The algorithms are Linear kernel SVM, RBF kernel SVM, LDA	Linear kernel SVM (C=0.1) with AEC Feature for Unconscious class	Accuracy ~ 85%
6	Avilov et al. (2021) (77)	Detection of intraoperative awareness during general anesthesia, especially	General	Propofol	-	Unavailable	22 (M: 10, F: 12)/Inria	Extracted features from EEG signal by CSP filters, Riemannian geometry, linear discriminant analysis	CSP+LDA, Minimal Distance to the Riemannian Mean, Tangent Space+LR, DeepConvNet, ShallowConvNet, EEGNet-2.32 & EEG-4.8	EEGNet-4.8 with 128 Electrodes	Accuracy: 94.5%, false-positive rate: 6.1%
7	Sook Ra et al. (2021) (78)	Develop a new DoA index for monitoring the DoA	-	Propofol	Consciousness, Light Anesthesia, Deep Anesthesia	Request	Not reported/University of Southern Queensland	Extracted features from EEG signals by entropy methods (SE and PE) and age, Weight, Height, Gender, Midazolam, Alfentanil, Propofol, Parecoxib, Fentanyl	SVM, DL Algorithm & NN, LR	LR	The Pearson correlation coefficient, RMSE, and execution time of LR were the best.

8	Sanz Perl et al. (2021) (92)	To develop whole-brain computational models to show that the stability of conscious states provides information complementary to their similarity to conscious wakefulness	-	Propofol	Sleep, Propofol Anesthesia & Post-Comatose Disorders of Consciousness	Part of the Dataset (Sleep data) is available	43 (M: 27, F: 16)/Medical School of the University of Liège	Extracted features from fMRI and EEG data & Gender, Age at Scan, Etiology, Days since Injury, Aud. Function, Vis. Function, Mot. Function, Oro/Ver Function, Communication, Aurosal, #CRS-R assessment	RF	RF	RF has good performance in this study.
9	Madanu et al. (2021) (3)	To extract features from EEG signals to predict DoA	General	Propofol	Anesthesia Deep, Anesthesia OK, Anesthesia Light	Unavailable	50/National Taiwan University Hospital	EEG data	CNN with 5, 6 & 10 layers, AlexNet, Pre-Trained VGG16, Pre-trained VGG19, Pre-trained InceptionRESV2	CNN with 10 layers	Accuracy: 83%
10	Fali Li et al. (2021) (79)	To detect the time onset at which patients lost their consciousness within the duration of (10, 61) after being injected with Propofol	General	Propofol	The Loss of Consciousness (LOC), Resting	Unavailable	30/Shanghai Sixth People's Hospital	Features extracted from EEG signals time-series by multi-channel cross-fuzzy entropy	Corresponding time-varying cross-fuzzy networks(C-FuzzyEn) time-varying coherence networks	C-FuzzyEn	The Long-Range Connectivity (LRC) of C-FuzzyEn is better than the LRC of COH (LRC is a parameter that measures the number of frontal-occipital connectivity)
11	Yuqin Li et al. (2021) (80)	To track the loss of consciousness and recovery of consciousness under General Anesthesia, using EEG signals	General	Propofol	Anesthesia Induction (i.e., LOC), Anesthesia Recovery (i.e., ROC), the Resting State (i.e., Resting)	Request	30 (M: 14, F: 16)/Shanghai Sixth People's Hospital	Three Feature sets, including Spatial Pattern of the Network features, Properties & SPN features + Properties	SVM	SVM	Accuracy: 95%, sensitivity: 93.33%, specificity: 96.67%
12	Tacke et al. (2020) (93)	Construct a combined electroencephalographic anesthesia index that predicts responsiveness in anesthetized patients.	General	Remifentanil, Sevoflurane, Propofol & Succinylcholine	-	Unavailable	39/not reported	EEG Parameters: Weighted Spectral Median Frequency, quotient of WSMF, Spectral Entropy, Hurst Exponent, Approximate Entropy, Lempel-Ziv Complexity, Permutation entropy Auditory Evoked Potentials Parameters: Wavelet Coefficients, Amplitudes and latencies of Wavelet Coefficients, Signal Energies based on Wavelet Coefficients, Maximum Amplitude of Retransformed AEPs, Variance of the Second Derivative of Wavelet Coefficients	SVM, NB Classifier, LR, MLP, Bayesian Net, J48	SVM	P_k: 0.935±0.11
13	Campbell et al. (2020) (91)	Identifying degrees of pathological unconsciousness in clinical patients under anesthesia via resting-state fMRI	General	Propofol, Remifentanil & Succinylcholine	Anesthesia-SHH: Awake, Light Sedation, Deep Sedation or General Anesthesia Anesthesia-WI: Wakefulness Baseline, Light Sedation, Deep Sedation, Recovery DOC: Healthy, Unresponsive Wakefulness Syndrome, Minimally Conscious State	Available	93/Anesthesia-SHH, Anesthesia-WI, DOC	Resting-State fMRI Feature Data: 32 features, including Amplitude of Low-Frequency Fluctuations (3), Within Network Functional Connectivity FC(8), Between Network Functional Connectivity FC(21)	SVM, Extra Trees, ANN	All	AUC>0.95
14	Ramaswamy et al. (2020) (81)	Estimate the depth of sedation via frontal EEG signals	-	Propofol, Dexmedetomidine, Sevofurane & Remifentanil	Awake, Sedated	Unavailable	66/Using a 16 channel Neuroscan® EEG monitor (Compumedics USA, Limited, Charlotte)	EEG Signals: Nonlinear energy operator, Activity, Mobility, Complexity, Root Mean Square Amplitude, Kurtosis, Skewness, Mean of Amplitude Modulation, Standard Deviation of AM, Skewness of AM, Kurtosis of AM, Burst Suppression ratio/min, Pδ=mean power in delta band, Pθ=mean power in theta band, Pα=mean power in alpha band, Pσ=mean power in spindle band, Pβ=power in beta band, PT=total spectral power, Pδ/PT, Pθ/PT, Pα/PT, Pσ/PT, Pβ/PT, Pδ/Pθ, Pα/Pθ, Pσ/Pθ, Pβ/Pθ, Pα/Pθ, Pσ/Pθ, Pβ/Pθ, Mean of Frequency Modulation, Standard Deviation of FM, and extra.	Elastic Net LR, SVM with Gaussian Kernel, RF, Ensemble Tree with Bagging	Ensemble Tree with Bagging	AUC: 0.88
15	Kashkooli et al. (2020) (82)	Design drug-specific models to improve the performance of automated anesthetic state monitors	General	Sevoflurane & Ketamine	Awake, Sedation, General Anesthesia	Unavailable	12 (M: 7, F: 5)/Waveguard system with a standard EEG cap(64 channels, ANT Neuro)	EEG data: Mean Power of Slow, Mean Power of Theta, Mean Power of Low-Beta, Mean Instantaneous Frequency, Kurtosis of Instantaneous Frequency, Hjorth Mobility, Permutation Entropy, Higuchi Fractal Dimension	KNN	KNN	F1-score: 94%

16	Lee et al. (2020) (83)	To identify brain states independent of the actual anesthetic concentration	-	Desflurane	-	Unavailable	7 Rats (M: 7, F: 0)/SmartBox (NeuroNexus Technologies)	Spike Rate, Local Variation, Total Number of Spikes, Longest Period Below Mean, Sample Entropy	Hierarchical Agglomerative Algorithm with Ward's Linkage Method	Hierarchical Agglomerative Algorithm with Ward's Linkage Method	-
17	Hayase et al. (2020) (84)	Improve anesthesia depth monitoring using the 20-Hz to 30-Hz hierarchical Poincaré analysis.	General, Local	Propofol, Sevofurane, Remifentanil & Fentanyl	Lighter Anesthesia, Deeper Anesthesia	Unavailable	30 (M: 16, F: 14)/Kyoto Chubu Medical Center	Poincaré-index20–30 Hz, Poincaré-index0.5–47 Hz, Electromyogram EMG70–110 Hz, Suppression Ratio	MLPNN	MLPNN	Correlation Coefficient: 0.87 RMSE: 7.09
18	Shalbaf et al. (2020) (85)	To assess the level of hypnosis with Sevoflurane	-	Sevoflurane	Awake, Light Anesthesia, General Anesthesia, Deep Anesthesia	Unavailable	17/not reported	EEG data: Frequency Index (Beta), Sample Entropy, Shannon Permutation Entropy, Detrended Fluctuation Analysis	SVM	SVM	Accuracy: 94.11%
19	Park et al. (2020) (86)	To present a real-time EEG-based DoA monitoring system	General	Sevoflurane & Propofol	-	Available	374/VitalDB constructed at Seoul National University Hospital	EEG data	ANESNET	ANESNET	MSE: 0.048 MAE: 0.05 Pearson Correlation Coefficient: 0.676 Concordance Correlation Coefficient: 0.566
20	Li et al. (2020) (87)	To monitor DoA	-	Sevoflurane	-	Request	20/Waikato Hospital in Hamilton	EEG data	LSTM and Sparse Denoising Autoencoder	LSTM and Sparse Denoising Autoencoder	P_k: 0.8556±0.0762
21	Ihalainen et al. (2021) (88)	Evaluate the evidence for the posterior hot zone theory of consciousness by modeling the relative contributions of three resting-state networks for Propofol-induced loss of consciousness.	General	Propofol	Behavioural Responsiveness, Sedation, Loss of Consciousness with Clinical Unconsciousness, Recovery of Consciousness	Request	10 (M: 4, F: 6)/Faculty of Medicine of the University of Liège	EEG data	Dynamic Causal Modelling (DCM) (Combination of 3 Networks: Default Mode Network (DMN), SAlience Network (SAN), Central Executive Network (CEN))	In Frontoparietal Connections: DMN In Frontal Connections: SAN In Parietal Connections: SAN All Connections: Combination of 3 Networks	AUC: 0.78, accuracy: 80%, mean posterior probabilities: 0.67, recall: 78%

Abbreviations: EEG, electroencephalogram; LDA, linear discriminant analysis; CNN, convolutional neural network; HMM, hidden markov model; BWP, bandwise power; PCA, principal component analysis; AUC, area under the curve; AA, anesthesia agent; BIS, bispectral index; ASA, American Society of Anesthesiology; DNN, deep neural network; DL, deep learning; LSTM, long short-term memory; DoA, depth of anesthesia; BMI, body mass index; LR, logistic regression; SVM, support vector machine; DT, decision tree; AEC, amplitude envelope correlation; wPLI, weighted phase lag index; FC, functional connectivity; RBF, radial basis function; CSP, common spatial pattern; NN, neural network; RMSE, root mean square error; RF, random forest; SPN, spatial pattern of the network; WSMF, weighted spectral median frequency; AEPs, auditory evoked potentials; NB, naïve bayes; MLP, multi-layer perceptron; P_k, prediction probability; DOC, disorders of consciousness; ANN, artificial neural network; AM, amplitude modulation; FM, frequency modulation; KNN, K-nearest neighbor; MLPNN, multi-layer perceptron neural network; MSE, mean square error; MAE, mean absolute error; DMN, default mode network; SAN, SAlience network.

Table 11. Category H: Control of Anesthesia Delivery

No.	Study	Goal	Type of Anesthesia	Induction Drug(s)	Dataset Availability	Number of Case/Dataset	Feature(s)	Algorithm(s)	Winner Algorithm	Winner Algorithm Performance	Interpretable?
1	Ingrande et al. (2020) (94)	To predict the dose of propofol anesthetic during surgery	General	Propofol	Unavailable	24 (M: 6, F: 18)/not reported	Gender, Age, Lean Body Weight, Total Body Weight, BMI, Cardiac Output	4-compartment Model, Recirculatory Model, GRU	GRU	MPE: 0.161 MSE: 20.83	No
2	Sharma et al. (2020) (98)	To improve the drug infusion to the automatic control of mean arterial blood pressure for maintaining the mean arterial pressure to 100 mmHg	General	Sodium & Nitroprusside	Unavailable	Not reported/not reported	-	Type-2 fuzzy logic, Cuckoo Algorithm (Optimization)	Type-2 fuzzy logic, Cuckoo Algorithm	Error ≈ 0	No
3	Miyaguchi et al. (2021) (95)	To determine whether Remifentanil will be given to the patient in the next n minutes.	General	Remifentanil	Unavailable	210 (M: 103, F: 107)/Okayama University Hospital	Static Features: Patient Information (age, weight, height, gender) & Dynamic Features: Vital Records (HR, SBP, DBP, MAP, RR, S_pO_2, $ETCO_2$) and Drug Records (Remifentanil flow)	SVM, LR, RF, ANN, LightGBM, LSTM	LSTM	Accuracy: 73%, sensitivity: 65%, specificity: 73%, precision: 2.3%, AUC: 0.75	Yes, using Shapley
4	Aiassa et al. (2021) (99)	Build a learning model for electrochemical sensing, which compensates for the fouling effect of propofol.	General	Propofol	Unavailable	480/not reported	Using 4 chemical feature sets including 1. ip, Ep, 2. ip, Ep, nmeas, 3. ip, Ep, Q, 4. ip, Ep, Q, nmeas	SVC with different kernels (Linear, polynomial, RBF, and sigmoid) & C=10 for all models	SVC (RBF)	Accuracy: 95%	No
5	Wei et al. (2021) (96)	To determine the appropriate dose of hyperbaric bupivacaine based on physical variables during cesarean section in the next 10 minutes	Neuraxial	Hyperbaric bupivacaine	Request	684/Ethical Committee of Jiaxing Maternity	Parturient demographic Features: Age, Weight, Height, Fundal height, Demographic Features: Vertebral column length, Abdominal girth, Fetal biparietal diameter, Fetal weight, Bupivacaine dosage	DT with different hyperparameters	DT (λ Value = 0.2)	MSE: 0.084	Yes, using Decision Tree Rules
6	Schamberg et al. (2022) (97)	To suggest the appropriate dose of anesthetic drug to automatically control the level of anesthesia during surgery	General	Propofol	Unavailable	9/not reported	Level of Unconsciousness (LoU) error, predicted effect-site concentration, LoU change, LoU target.	DRL	DRL	Performance Error: 0.011±0.005	Yes, using Shapley additive explanations

Abbreviations: BMI, Body Mass Index; GRU, Gated Recurrent Unit; MPE, mean percentage error; MSE, mean square error; HR, heart rate; SBP, systolic blood pressure; DBP, diastolic blood pressure; MAP, mean arterial pressure; RR, respiratory rate; S_pO_2, peripheral oxygen saturation; $ETCO_2$, end-tidal carbon dioxide; SVM, support vector machine; LR, logistic regression; RF, random forest; ANN, artificial neural network; LightGBM, light gradient boosting machine; LSTM, long short-term memory; AUC, area under the curve; SVC, support vector classifier; RBF, radial basis function; DT, decision tree; LoU, level of unconsciousness; DRL, deep reinforcement learning.