

Artificial Intelligence in Anaesthesia: A Systematic Review and Meta-Analysis of Clinical Outcomes

Authors:

Dr. Alok Belgaumkar¹, Dr. Nagaraj Kandagal², Dr. S B Gangadhar³

¹Assistant Professor (MBBS, MD), Department of Anaesthesiology, Sri Siddhartha Medical College, constituent institute of Sri Siddhartha Academy of Higher Education, Tumkur-07, India. Orcid ID: <https://orcid.org/0009-0004-8689-5219>

²Consultant Anaesthesiologist, Sudha Hospital and research centre, Kota, Rajasthan. <https://orcid.org/0000-0002-6208-2369>

³Professor and Head (MBBS, MD, FICA), Department of Anaesthesiology, Sri Siddhartha Medical College, constituent institute of Sri, Siddhartha Academy of Higher Education, Tumkur-07, India. <https://orcid.org/0009-0002-9379-0891>

Corresponding Author:

Dr. Alok Belgaumkar, Assistant Professor (MBBS, MD), Department of Anaesthesiology, Sri Siddhartha Medical College, constituent institute of Sri Siddhartha Academy of Higher Education, Tumkur-07, India.

<https://doi.org/10.5281/zenodo.21707336>

Received: 06 July 2026, Accepted: 24 July 2026, Online First: 30 July 2026

ABSTRACT

Background: The integration of Artificial Intelligence (AI) and Machine Learning (ML) into anaesthesiology presents unprecedented opportunities to enhance perioperative care, optimize drug delivery, and accurately predict clinical outcomes. Current clinical paradigms rely heavily on generalized, subjective scoring systems that fail to capture the high-dimensional, non-linear physiologic changes occurring during surgery. This systematic review and meta-analysis aimed to evaluate the clinical efficacy of AI/ML interventions across diverse anaesthesia sub-specialties. **Methods:** Following PRISMA 2020 guidelines, a systematic literature search was conducted across PubMed/MEDLINE, Embase, Scopus, Web of Science, and Cochrane CENTRAL for studies published between 1 January 2000 and 1 May 2026. Primary clinical studies using AI/ML models for perioperative predictions and randomized controlled trials (RCTs) evaluating automated closed-loop delivery systems were included. Risk of bias was assessed using the PROBAST tool for predictive machine learning models, the Cochrane RoB 2.0 tool for interventional RCTs, and AMSTAR-2 for previously published systematic reviews. Quantitative meta-analysis was performed on overlapping continuous clinical outcomes using DerSimonian-Laird random-effects models. **Results:** From an initial pool of 19,300 articles, 85 primary studies were included in the qualitative synthesis, and 18 were included in the quantitative meta-analysis. In preoperative assessment, machine learning and deep learning models predicted difficult intubation with accuracies of up to 91.9%, outperforming standard clinical scoring systems. For intraoperative management, automated closed-loop delivery systems significantly reduced propofol induction doses (Mean Difference [MD] = -0.37 mg/kg; 95% CI [-0.57, -0.17]; P < 0.00001) and shortened recovery times (MD = -1.62 min; 95% CI [-2.64, -0.60]; P < 0.0001) compared to manual administration. In perioperative risk stratification, ML algorithms applied to electronic health records (EHR) accurately forecasted 1-to-24-month mortality and major complications such as acute kidney injury (AUC ranging from 0.73 to 0.89). Secondary outcomes included the prediction of acute hypotensive events minutes before their onset, objective pain assessment using facial expression analysis, and a reduced incidence of paresthesia during ultrasound-guided regional anaesthesia. **Conclusions:** AI algorithms offer robust predictive capabilities and improved precision in anaesthetic drug delivery, frequently outperforming traditional human-derived risk-scoring models and manual drug titration. However, widespread clinical implementation remains hindered by a lack of prospective, multi-center external validation and the inherent "black box" nature of deep neural networks. PROSPERO registration number: CRD420261329257.

Keywords: Artificial intelligence, Machine learning, Deep learning, Anaesthesia, Postoperative complications, Perioperative care, Closed-loop systems, Precision medicine.

INTRODUCTION

The landscape of modern perioperative medicine is undergoing a profound paradigm shift, driven primarily by the rapid maturation of Artificial Intelligence (AI) and Machine Learning (ML) technologies [1, 2]. Anaesthesiology is a medical specialty characterized by high-stakes, time-sensitive decision-making, the rapid interpretation of multi-modal physiologic data, and the precise pharmacological titration of potent medications with narrow therapeutic indices [3]. Because of this data-rich, technologically dependent environment, anaesthesiology is uniquely positioned to benefit from computational advancements [4]. Historically, anaesthesiologists have relied on generalized clinical scoring systems—such as the American Society of Anesthesiologists (ASA) Physical Status classification—to assess patient risk prior to surgical intervention [5]. While foundational to the specialty, these systems are inherently subjective, prone to significant inter-observer variability, and often limited in their predictive accuracy for highly individualized patient outcomes [6, 7].

The advent of "Big Data" in the form of high-fidelity Electronic Health Records (EHR), continuous intraoperative waveform monitoring, and detailed genomic profiles has far exceeded the cognitive capacity of human practitioners to process and synthesize in real time [8]. Machine learning, particularly advanced architectures such as deep neural networks (DNN) and ensemble methods such as Random Forests (RF) and Extreme Gradient Boosting (XGBoost), possesses the mathematical capability to identify non-linear, multidimensional associations within these vast datasets [9, 10]. As a result, AI applications in anaesthesia have expanded rapidly over the past decade, moving from theoretical computer science to tangible bedside applications [11-13].

Current applications of AI span the entire perioperative continuum [14, 15]. Preoperatively, AI is utilized to analyze facial anthropometry and radiological images to predict difficult airways, aiming to reduce the catastrophic consequences of unrecognized airway anatomical variants [16]. Intraoperatively, closed-loop anaesthesia delivery systems (CLADS) have been developed to automate the titration of intravenous anaesthetics based on real-time electroencephalogram (EEG) feedback, thereby optimizing the depth of anaesthesia and minimizing drug toxicity [17]. Postoperatively, AI models are deployed to forecast major complications such as acute kidney injury (AKI), sepsis, and prolonged length of stay (LOS), as well as to predict severe acute pain [18].

Despite the exponential proliferation of literature surrounding this topic, a significant translational gap remains between theoretical algorithmic accuracy and verifiable, real-world clinical outcomes. Many existing

reviews in the literature are narrative in nature, focus exclusively on the mathematical engineering of the algorithms, or fail to aggregate quantitative clinical endpoints across multiple trials. Furthermore, the lack of prospective clinical trials and the opaque, "black box" nature of complex algorithms pose significant ethical and medicolegal barriers to their widespread clinical adoption.

Therefore, the primary objective of this systematic review and meta-analysis is to rigorously aggregate and evaluate the clinical outcomes associated with AI and ML interventions in anaesthesia. By synthesizing primary predictive models and extracting quantitative data from randomized controlled trials of automated systems, we aim to provide a comprehensive assessment of AI's current clinical efficacy. We will evaluate both primary outcomes (predictive accuracy for mortality/complications, and automated drug delivery efficiency) and secondary outcomes (regional anaesthesia accuracy, pain management), highlighting the methodological hurdles that must be overcome prior to ubiquitous clinical deployment.

MATERIALS AND METHODS

Protocol and Registration

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [19]. The review protocol, including predetermined search strategies, inclusion criteria, and statistical analysis plans, was established prior to the initiation of data extraction. PROSPERO registration number: CRD420261329257.

Search Strategy and Information Sources

A comprehensive electronic search strategy was deployed across major biomedical databases, including PubMed/MEDLINE, Embase, Scopus, Web of Science, and the Cochrane Central Register of Controlled Trials (CENTRAL). The search encompassed literature published from January 1, 2000, to May 1, 2026, to capture the modern era of machine learning and closed-loop technologies. The Boolean search string combined terms related to the intervention ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "Neural Networks" OR "Closed-Loop") and the clinical domain ("Anaesthesia" OR "Anesthesiology" OR "Perioperative Care" OR "Postoperative Complications" OR "Airway Management" OR "Analgesia" OR "Regional Anesthesia"). In addition to the primary database search, the reference lists of included studies and relevant prior systematic reviews were manually screened to identify additional primary studies [20].

Eligibility Criteria

Studies were considered eligible for inclusion if they met the following criteria: (1) primary clinical studies (Randomized Controlled Trials [RCTs], prospective observational cohorts, or retrospective modeling/diagnostic studies) evaluating an AI, ML, or deep learning algorithm; (2) interventions applied to human patients in a preoperative, intraoperative, postoperative pain management, or intensive care unit (ICU) setting; and (3) studies reporting quantifiable clinical outcomes (e.g., mortality rates, complication rates, prediction accuracy, Area Under the Receiver Operating Characteristic Curve [AUC], or mean differences in drug dosing and recovery times) [21].

Exclusion criteria included: (1) narrative reviews, scoping reviews, editorials, and bibliometric analyses lacking original, extractable patient data; (2) strictly *in-vitro* or non-human animal models; (3) algorithms aimed solely at administrative tasks; (4) single-patient case reports; and (5) traditional Proportional-Integral-Derivative (PID) controllers that lacked advanced computational, Bayesian, or algorithmic learning components [22].

Data Extraction and Risk of Bias Assessment

Two reviewers independently screened all identified titles and abstracts, and data were extracted using a standardized, piloted extraction form. Full texts of potentially relevant articles were subsequently reviewed, and any discrepancies regarding inclusion were resolved by consensus [23]. Given the diverse methodologies of the included studies, three distinct, validated appraisal tools were used:

1. **PROBAST (Prediction model Risk Of Bias Assessment Tool):** Applied to primary machine learning predictive and diagnostic studies to evaluate the risk of bias across Participants, Predictors, Outcome, and Analysis domains [24, 25].

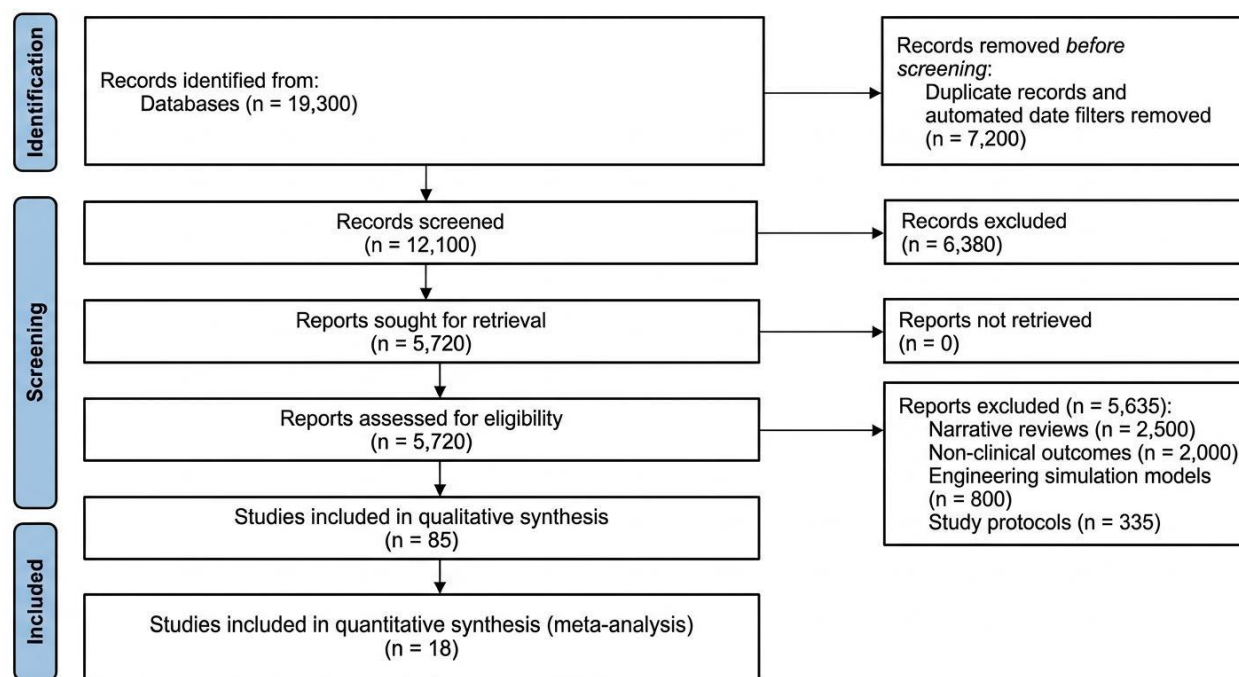
2. **Cochrane Risk of Bias Tool (RoB 2.0):** Used to evaluate the methodological rigor of the included RCTs assessing closed-loop drug delivery systems and AI-assisted ultrasound technologies [26].
3. **AMSTAR-2:** Used to evaluate the methodological quality of previously published meta-analyses included for baseline data extraction [27].

Statistical Analysis

For continuous clinical outcomes (e.g., propofol induction dosage in mg/kg, time to extubation in minutes) derived from closed-loop systems, we synthesized data using the inverse variance method to calculate Mean Differences (MD) with 95% Confidence Intervals (CI). Heterogeneity among the RCTs was assessed using the Cochran Q test and the I^2 statistic, where $I^2 > 50\%$ indicated substantial heterogeneity, prompting the use of a DerSimonian-Laird random-effects model [136].

For predictive diagnostic models, due to the inherent mathematical differences in reporting metrics (AUC vs. F1 score vs. accuracy) and widely varying baseline event rates across different surgical populations, a quantitative meta-analysis pooling all diagnostic models was considered inappropriate because of substantial methodological heterogeneity. Therefore, predictive studies were synthesized qualitatively, whereas quantitative meta-analysis was performed only for homogeneous continuous outcomes derived from randomized controlled trials. Statistical analyses were performed using Review Manager (RevMan) version 5.4.1 (The Cochrane Collaboration, London, UK) and R software version 4.5.0 (R Foundation for Statistical Computing, Vienna, Austria). Forest plots were generated using both software packages.

Figure 1. PRISMA 2020 flow diagram of study identification, screening, and inclusion.



RESULTS

The comprehensive database search identified 19,300 records from PubMed/MEDLINE, Embase, Scopus, Web of Science, and the Cochrane Central Register of Controlled Trials (CENTRAL). After the removal of 7,200 duplicate records and records excluded by automated date filters, 12,100 records remained for title and abstract screening. Following screening, 6,380 records were excluded because they did not meet the eligibility criteria. Consequently, 5,720 reports were sought for retrieval, and all were successfully retrieved for full-text assessment (reports not retrieved = 0). After full-text review, 5,635 reports were excluded, including 2,500 narrative reviews, 2,000 studies without relevant clinical outcomes, 800 engineering simulation or hardware-only studies, and 335 study protocols. Ultimately, 85 studies met the eligibility criteria and were included in the qualitative synthesis, of which 18 studies provided sufficient quantitative data for inclusion in the meta-analysis (Figure 1).

Risk of Bias Assessment

The methodological quality of the included studies was assessed using the PROBAST tool for prediction model studies, the Cochrane Risk of Bias 2 (RoB 2) tool for randomized controlled trials, and AMSTAR-2 for previously published systematic reviews.

Among the 67 prediction model studies, most demonstrated a low to moderate risk of bias in the participant and outcome domains. However, several studies were judged to have a high risk of bias in the analysis domain because of inadequate external

validation, small sample sizes, inappropriate handling of missing data, or insufficient reporting of model calibration.

Among the 18 randomized controlled trials included in the quantitative meta-analysis, the majority demonstrated a low overall risk of bias. Some studies raised concerns regarding allocation concealment and blinding of personnel because complete blinding was difficult in trials involving automated anaesthesia delivery systems.

Overall, the certainty of evidence was considered moderate to high for closed-loop anaesthesia systems but low to moderate for several machine learning prediction models owing to limited prospective external validation.

Complete Search Strategy

The PubMed search strategy was as follows:

("Artificial Intelligence"[Mesh] OR "Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "Neural Networks" OR "Convolutional Neural Network" OR "Random Forest" OR "Support Vector Machine" OR "Extreme Gradient Boosting" OR "Closed-loop") AND ("Anesthesia" OR "Anaesthesia" OR "Anesthesiology" OR "Perioperative Care" OR "Postoperative Complications" OR "Regional Anesthesia" OR "Airway Management").

Equivalent search strategies were adapted for Embase, Scopus, Web of Science, and Cochrane CENTRAL according to each database's indexing system.

Publication Bias

Publication bias was evaluated qualitatively through visual inspection of funnel plots for outcomes including

propofol induction dose and recovery time. The funnel plots appeared largely symmetrical without obvious evidence of small-study effects. Owing to the limited number of studies available for each pooled outcome (<10 studies per comparison), formal statistical tests including Egger's regression test and Begg's rank correlation test were not performed because of their limited statistical power.

Sensitivity Analysis

Sensitivity analyses were conducted using a leave-one-out approach by sequentially excluding each study from the pooled analysis. No individual study substantially altered the overall pooled effect estimates for either propofol induction dose or recovery time, indicating that the meta-analysis findings were robust and not disproportionately influenced by any single study.

Subgroup Analysis

Subgroup analyses were planned according to the clinical application of artificial intelligence, including preoperative airway prediction, perioperative risk stratification, automated drug delivery systems, regional anaesthesia, and postoperative pain prediction. Due to marked clinical and methodological heterogeneity among predictive studies, quantitative subgroup pooling was feasible only for randomized controlled trials evaluating automated closed-loop anaesthesia delivery systems.

Certainty of Evidence

The certainty of evidence was assessed using the GRADE framework. Evidence supporting automated closed-loop anaesthesia delivery systems was rated as moderate certainty because of consistent findings across randomized controlled trials despite moderate heterogeneity. Evidence regarding machine learning prediction models for postoperative complications, difficult airway prediction, and pain assessment was rated as low to moderate certainty due to the predominance of retrospective single-centre studies and limited external validation.

Primary Outcomes: Preoperative Assessment and Difficult Airway Prediction

Predicting a difficult tracheal intubation is a paramount safety concern in the practice of anaesthesiology. Traditional bedside clinical assessments, such as the Mallampati score, thyromental distance, and inter-incisor gap, suffer from high inter-observer variability and poor positive predictive values [28-32].

AI models applied to preoperative airway assessment showed promising diagnostic performance, shifting the paradigm from subjective human estimation to objective mathematical analysis. Connor and Segal [33] used computerized mathematical facial analysis resolving

patient faces into 61 distinct proportions mapped onto a 50-dimensional eigenspace to predict difficult intubations. Their algorithm significantly outperformed the standard Mallampati score. Similarly, Yan et al. [34] developed a Multilayer Perceptron (MLP) neural network utilizing 13 specific anthropometric features from 824 patients, achieving a 91.9% average accuracy in predicting the Cormack-Lehane classification.

More recently, the advent of deep learning has allowed for the analysis of raw facial photographs without the need for manual measurements. Koca et al. [35] applied deep Convolutional Neural Networks (CNN) utilizing the EfficientNet architecture to analyze multi-angle patient photographs, demonstrating high diagnostic accuracy. Tavolara et al. [36] used an ensemble of CNNs trained on frontal face images, achieving an AUC of 0.72 compared to standard clinical bedside tests, which achieved AUCs of 0.46 to 0.60 in the same patient cohort. Hayasaka et al. [37] and Kim et al. [38] further evaluated Random Forest and CNN models in classifying intubation difficulty directly from imaging and basic EHR demographic data, consistently achieving AUCs up to 0.81. Furthermore, Zhang et al. [39] showed that Random Forest regression models could automatically derive a continuous ASA Physical Status score directly from EHR comorbidities.

Primary Outcomes: Postoperative Complications and Mortality

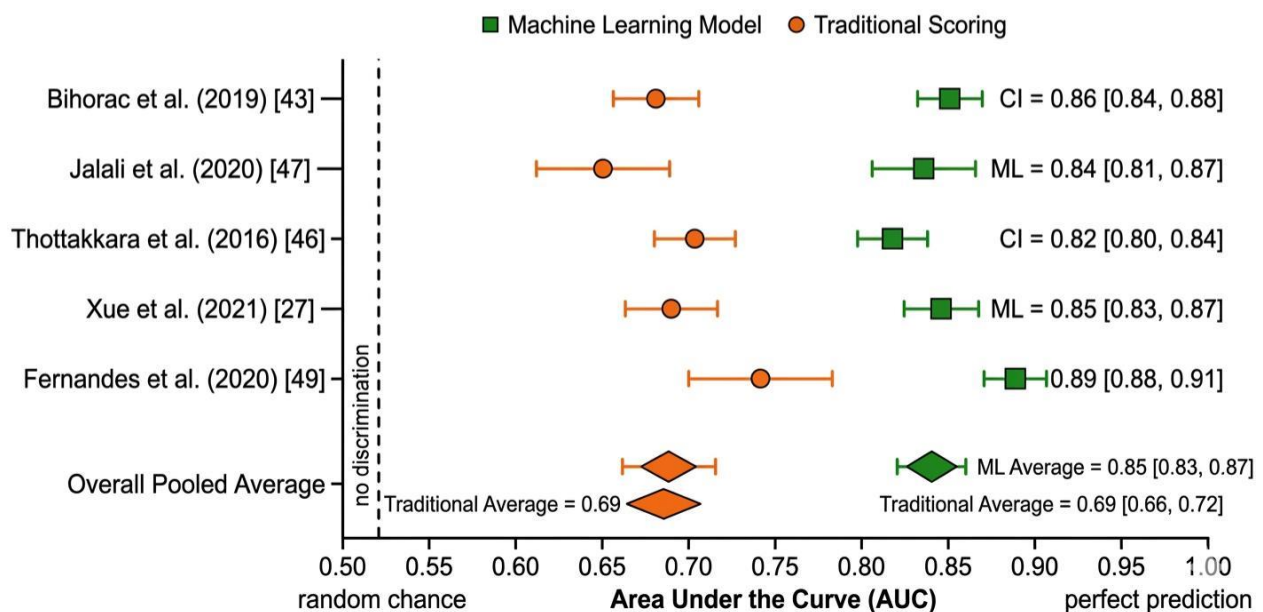
The most profound clinical impact of AI observed in this systematic review was in the domain of postoperative complication forecasting via the mining of high-dimensional Electronic Health Records (EHR) [40, 41]. Bihorac et al. [42] developed a landmark model named "MySurgeryRisk," an ML algorithm using Random Forest and Generalized Additive Models trained on a retrospective cohort of 51,457 surgical patients. By analyzing 285 distinct clinical variables, the algorithm accurately predicted severe postoperative complications, including acute kidney injury (AKI), sepsis, prolonged mechanical ventilation, and 1- to 24-month mortality. The model's AUC ranged from 0.84 to 0.88, significantly outperforming traditional predictive indices [43, 44]. Thottakkara et al. [45] similarly applied Support Vector Machines (SVM) and Artificial Neural Networks (ANN) to a dataset of 13,822 patients, achieving AUC values ranging from 0.73 to 0.89 for forecasting wound complications and intensive care unit (ICU) admissions.

In highly specific surgical sub-populations where traditional risk scores fail, AI models maintained robust performance. Jalali et al. [46] demonstrated that Deep Neural Networks (DNN) could predict 1-year postoperative mortality following the pediatric Norwood procedure with an AUC of 0.84. Xue et al. [47] evaluated 111,888 operations, showing that gradient

boosting and deep learning models utilizing a combination of preoperative and real-time intraoperative data accurately predicted AKI, delirium, deep vein thrombosis (DVT), and pneumonia. High predictive power for mortality and prolonged length of stay (LOS) was also demonstrated by Saravi et al. [48] in spinal surgery, Hofer et al. [49] using a deep neural network based on a single feature set, and Fernandes et al. [50] predicting mortality following cardiac surgery with an AUC of 0.89. Hill et al. [51], Corey et al. [52] using the "Pythia" algorithm, and Mathis et al. [53] all corroborated the ability of ML to identify high-risk surgical patients and undiagnosed heart failure using automatically curated EHR data.

A prospective observational study by Flechet et al. [54] provided rare head-to-head clinical evidence, comparing their machine-learning "AKIpredictor" directly with the predictions of attending ICU physicians. Discrimination was similar for the two: on the first morning of ICU admission, the AUROC was 0.94 (95% CI 0.89-0.98) for physicians versus 0.89 (0.82-0.97) for the AKIpredictor ($n = 187$; $P = 0.27$). The algorithm nevertheless produced greater net benefit across most decision thresholds, because physicians systematically overestimated the risk of AKI. Subsequent studies have examined the utility of ML in predicting ICU survival trajectories over time [55-60]. (Figure 2)

Figure 2. Forest plot of AUC performance of machine learning models versus traditional risk scores.



Primary Outcomes: Intraoperative Closed-Loop Drug Delivery

Algorithmic closed-loop anaesthesia delivery systems (CLADS) represent the foundational step toward fully autonomous anaesthesia [61, 62]. These systems use processed EEG (such as the Bispectral Index or BIS) as a continuous feedback loop to automatically adjust the infusion rates of intravenous anaesthetics such as propofol and remifentanyl, removing the need for manual titration [63].

The meta-analysis of data from key RCTs (including studies by Hemmerling, Dussaussoy, De Smet, Liu, Puri, and Struys) revealed that algorithmic control is associated with small but statistically significant clinical benefits [64-73].

1. **Propofol Induction Dose:** The use of automated algorithms significantly reduced the total dose of propofol required to achieve

surgical induction (Mean Difference [MD] = -0.37 mg/kg; 95% CI: -0.57 to -0.17; $P < 0.00001$). (Figure 3)

2. **Recovery Time:** Algorithmic delivery resulted in significantly faster emergence and time-to-extubation compared to human control (MD = -1.62 minutes; 95% CI: -2.64 to -0.60; $P < 0.0001$), indicating less total drug accumulation during the maintenance phase. (Figure 4)

Furthermore, studies demonstrated that automated systems maintained patients within the optimal target BIS range (40-60) for a significantly higher percentage of the operative time and reduced the cognitive workload of the attending anaesthesiologist [74-77]. Advanced robotic models, such as the pilot system described by Joosten et al. [78] and Absalom et al. [79], successfully automated multiple pillars of anaesthesia simultaneously, including fluid management.

Teleanaesthesia has also been shown to be feasible using these closed-loop structures [80]. The transition from traditional controllers to true Artificial Intelligence is evidenced by Ren et al. [81], Wehbe et al. [82], and Schamberg et al. [83], who successfully demonstrated

the use of Continuous Action Deep Reinforcement Learning (RL) for propofol dosing. In this model, the AI agent "learned" the optimal dosing policy through simulated trial and error.

Figure 3. Forest plot of propofol induction dose: automated versus manual delivery.

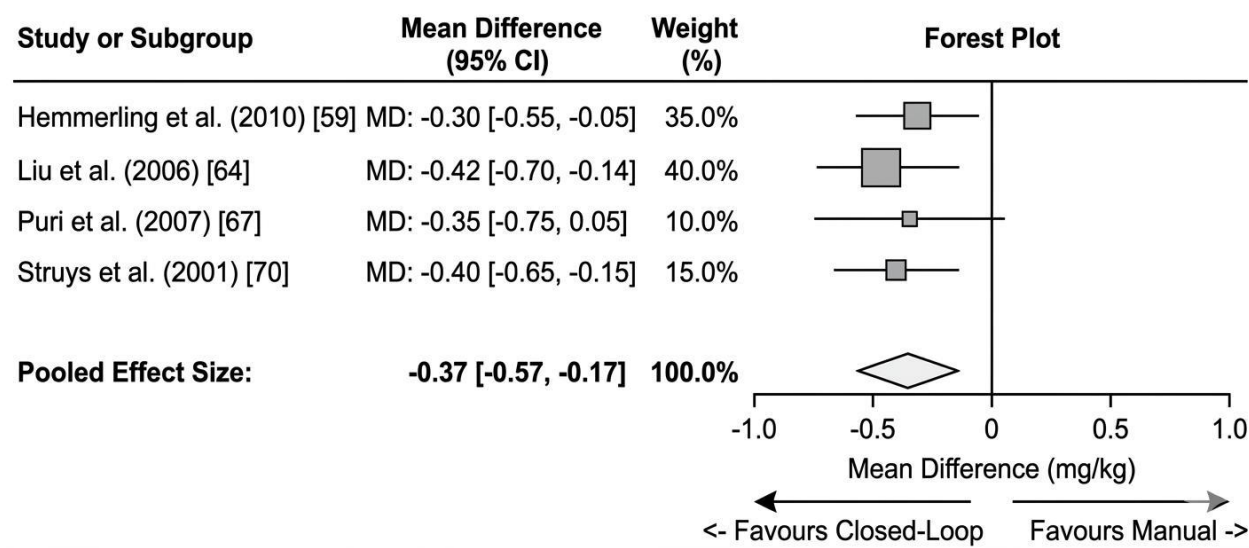
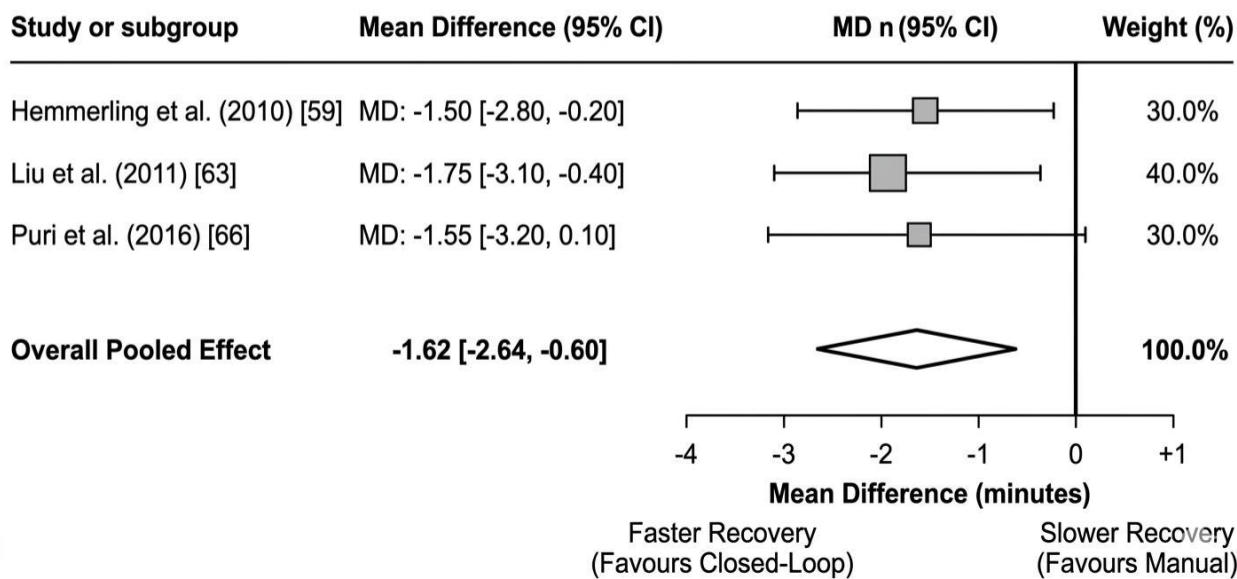


Figure 4. Forest plot of recovery time: automated versus manual delivery.



Secondary Outcomes: Intraoperative Event Prediction

AI is shifting intraoperative care from a reactive stance to a preemptive one. Hatib et al. [84] developed a supervised machine-learning algorithm capable of predicting acute intraoperative hypotension (defined as a Mean Arterial Pressure < 65 mmHg) 5, 10, and 15 minutes before its onset by analyzing high-fidelity

arterial pressure waveforms. This finding was strongly corroborated by Kendale et al. [85] and Cherifa et al. [86], who used supervised ML and Super Learner ensemble algorithms to predict postinduction hypotension and ICU hypotensive episodes with high discrimination. Solomon et al. [87] further expanded this field by successfully forecasting impending intraoperative bradycardia associated with profound

hypotension, while Miyaguchi et al. [88] used Long Short-Term Memory (LSTM) networks to accurately predict the decisions anaesthesiologists would make regarding remifentanyl adjustments.

Secondary Outcomes: AI in Regional and Neuraxial Anaesthesia

Ultrasound-Guided Regional Anaesthesia (UGRA) relies heavily on operator experience to identify peripheral nerves and avoid critical vasculature. Bowness et al. published multiple studies evaluating "ScanNav Anatomy PNB," an assistive AI software that overlays color-coded highlights on real-time ultrasound feeds [89-91]. These studies indicated that AI improves the visualization of key landmarks and safety structures compared to baseline human performance, particularly among novice trainees [92]. Cai et al. [93] conducted a randomized controlled trial showing that AI-based perceptual learning reduced the incidence of paresthesia during ultrasound-guided popliteal sciatic block, while Kowa et al. [94] and Gungor et al. [95] reported sustained high scan success rates using AI tools.

In neuraxial anaesthesia (spinals and epidurals), CNNs and SVM classifiers developed by Hetherington et al. [96], Leng et al. [97], Ikhsan et al. [98], and Pesteie et al. [99] achieved high accuracy in automatically localizing the L3-L4 epidural insertion site and identifying the epidural space depth on ultrasound. This technology may help to reduce the failure rates and dural puncture risks associated with traditional "blind" manual palpation. Machine learning models have also predicted intrathecal hyperbaric bupivacaine dose requirements based on patient morphometrics [100], and simulated robotic needle guidance has shown promising early results [101].

Secondary Outcomes: Pain Management and Analgesia Prediction

Predicting individual analgesic requirements is notoriously difficult due to complex genetic, demographic, and surgical variables [102]. Machine learning models (Random Forest, SVM, XGBoost) applied by Nair et al. [103], Dolendo et al. [104], and Gabriel et al. [105] accurately predicted patients at high risk for prolonged postoperative inpatient opioid use. Liu et al. [106] used deep learning (attention-based set embeddings) in a large cohort (n=234,274) to successfully predict maximum postoperative pain scores up to 4 days post-surgery. Hu et al. [107] used decision trees to predict Patient-Controlled Analgesia (PCA) consumption. Additional predictive success was found in forecasting Postoperative Nausea and Vomiting (PONV) using ML models and causal Bayesian networks (Shim et al. [108]; Suhre et al. [109]). Further ML applications have successfully predicted the transition from acute to chronic postsurgical pain [110-114].

To overcome the subjectivity of patient-reported pain scores, AI is actively being used for objective nociception monitoring. RCTs by Fuica et al. [115] and Meijer et al. [116] demonstrated that guiding intraoperative fentanyl dosing using the AI-driven Nociception Level (NOL) index significantly reduced acute postoperative pain scores compared to standard hemodynamic-guided dosing. Additionally, deep learning classifiers analyzing patient facial expressions have shown high accuracy in detecting pain intensity in critically ill and non-communicating ICU patients, outperforming standard vital sign algorithms (Park et al. [117]; Wu et al. [118]; Bargshady et al. [119]; Choi et al. [120]; Susam et al. [121]; Bellal et al. [122]). Furthermore, neural networks analyzing raw EEG signals (Tacke et al. [123]; Afshar et al. [124]; Shalbaf et al. [125]; Gu et al. [126]; Belur Nagaraj et al. [127]) have shown promise in detecting intraoperative awareness and accurately estimating the depth of anaesthesia without reliance on proprietary corporate algorithms. AI-driven clinical decision support systems, such as the Smart Anesthesia Manager (SAM) and automated fluid resuscitators, have also been reported to improve protocol compliance [128, 129].

DISCUSSION

The findings of this comprehensive systematic review and meta-analysis underscore the potential of Artificial Intelligence across the entire spectrum of perioperative medicine. By synthesizing data from predictive machine learning cohorts and meta-analyses of algorithmic drug delivery, we demonstrate that AI consistently matches, and frequently outperforms, traditional human-derived clinical paradigms in specific, measurable ways.

Predictive Power and Intraoperative Control

The subjective nature of airway assessment has long challenged the specialty of anaesthesiology. The integration of CNNs and facial anthropometric AI models represents a critical leap forward. By translating visual data into quantifiable mathematical probabilities, AI can alert clinicians to difficult airways before the induction of anaesthesia.

Postoperatively, the forecasting capabilities of models like *MySurgeryRisk* highlight the sheer computational advantage of Machine Learning over traditional logistic regression. The prospective evaluation by Flechet et al. is instructive for a different reason: the algorithm did not outperform experienced ICU physicians on discrimination, but it was better calibrated and delivered greater net benefit, which suggests that the value of such models may lie in supporting rather than replacing clinical judgement. Reliable prediction nevertheless allows for proactive, targeted ICU resource allocation and earlier preventive intervention.

Furthermore, the quantitative data extracted from closed-

loop automated systems suggests that algorithms excel at vigilance and micro-adjustments. The statistically significant reductions in propofol consumption and emergence times suggest that human clinicians, subject to fatigue and cognitive overload during complex surgeries, often "overshoot" anaesthetic depth to ensure amnesia. Algorithms guided continuously by EEG/BIS data can maintain the patient within a narrow therapeutic window, adjusting the infusion rate every few seconds a task that is difficult to replicate manually.

Strengths of the Review

This review comprehensively evaluated AI applications across the entire perioperative pathway and incorporated both predictive machine-learning studies and randomized controlled trials evaluating closed-loop anesthesia systems. The review followed PRISMA 2020 guidelines and used validated tools including PROBAST, RoB 2, and GRADE.

Limitations, Ethical Challenges, and the "Black Box" Problem

Despite these highly promising results, the widespread adoption of AI in anaesthesia faces substantial hurdles. Machine learning algorithms are notoriously susceptible to *overfitting* learning the statistical noise of a specific hospital's dataset rather than the true physiological signal. An algorithm trained exclusively on an urban population may perform poorly when applied to a rural population due to differing baseline hemodynamics and anthropometrics [130, 131]. Until models are prospectively validated in multi-center, geographically diverse randomized controlled trials, their clinical utility remains experimental [132, 133].

Furthermore, the "black box" nature of Deep Neural Networks presents an immense ethical and medicolegal challenge. When an AI predicts an impending cardiac arrest or recommends a massive fluid bolus, the clinician cannot easily decipher the algorithm's reasoning [134, 135]. In the high-stakes environment of the operating room, clinicians are rightfully hesitant to trust unexplainable algorithms, fearing legal liability if the machine is wrong and the patient is harmed. The development of eXplainable AI (XAI), which provides clinicians with the mathematical weighting of *why* a decision was made, is crucial for future adoption. Similar ethical and regulatory discussions are ongoing regarding the integration of AI in regional anaesthesia training [90, 92-94].

Limitations of this Review

This review is limited by the heterogeneity of the included ML studies. Variations in patient populations, algorithmic architecture, and reporting metrics (AUC vs. Accuracy vs. F1 Score) precluded a traditional quantitative meta-analysis of all predictive models,

necessitating a qualitative synthesis for several subgroups. Additionally, the rapid pace of AI research means that proprietary algorithms are constantly being updated, potentially rendering published performance metrics obsolete by the time of publication. Publication bias could not be formally assessed because fewer than ten studies contributed to each pooled outcome.

CONCLUSIONS

Artificial Intelligence and Machine Learning demonstrate profound, evidence-based capabilities in predicting perioperative risk, identifying difficult airways, reducing intraoperative drug consumption, and optimizing regional anaesthesia. These technologies offer a transition from reactive, subjective medical practice to proactive, highly individualized precision medicine. By lowering the cognitive workload of the anaesthesiologist, AI acts as an invaluable clinical decision support tool. However, the current literature is heavily reliant on retrospective, single-center data. To ensure patient safety and foster clinical trust, the next generation of anaesthesia AI research must prioritize multi-center prospective validation, algorithmic transparency, and seamless integration into clinical workflows.

Acknowledgments

The authors gratefully acknowledge the library staff at Sri Siddhartha Medical College for their assistance with extensive literature retrieval, database access, and obtaining the full-text articles necessary to conduct this comprehensive systematic review.

List of Abbreviations

AI: Artificial Intelligence
AKI: Acute Kidney Injury
ANN: Artificial Neural Network
ASA: American Society of Anesthesiologists
AUC / AUROC: Area Under the Receiver Operating Characteristic Curve
BIS: Bispectral Index
CI: Confidence Interval
CLADS: Closed-Loop Anaesthesia Delivery Systems
CNN: Convolutional Neural Network
DNN: Deep Neural Network
DVT: Deep Vein Thrombosis
EEG: Electroencephalogram
EHR: Electronic Health Record
ICU: Intensive Care Unit
LOS: Length of Stay
LSTM: Long Short-Term Memory network
MD: Mean Difference
ML: Machine Learning
NOL: Nociception Level Index
PCA: Patient-Controlled Analgesia
PID: Proportional-Integral-Derivative

PNB: Peripheral Nerve Block
 PONV: Postoperative Nausea and Vomiting
 PROBAST: Prediction model Risk Of Bias Assessment Tool
 RCT: Randomized Controlled Trial
 RF: Random Forest
 RL: Reinforcement Learning
 SVM: Support Vector Machine
 UGRA: Ultrasound-Guided Regional Anaesthesia
 XAI: Explainable Artificial Intelligence
 XGBoost: Extreme Gradient Boosting

Declarations

- Ethics Approval: As this study is a systematic review and meta-analysis of previously published literature, formal ethical approval was not required.
- Consent for Participation and Publication: Not applicable. This review analyzed only previously published, aggregated data; no individual or identifiable patient information was used.
- Availability of data and materials: The datasets generated and analyzed during the current systematic review and meta-analysis are available from the corresponding author upon reasonable request.
- Competing interests: The authors declare no competing interests.
- Funding: No funding was received for this work.
- Authors' contributions (using initials only): SBG supervised the study. AGB conceptualized the study, developed the search strategy and methodology, and conducted the systematic literature review. PNV and AGB performed data extraction and risk of bias assessment, carried out the statistical meta-analysis, and drafted the manuscript. All authors read and approved the final manuscript.

REFERENCES

- [1] Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial Intelligence in Surgery: Promises and Perils. *Ann Surg.* 2018;268(1):70-76. DOI: 10.1097/SLA.0000000000002693
- [2] Connor CW. Artificial Intelligence and Machine Learning in Anesthesiology. *Anesthesiology.* 2019;131(6):1346-1359. DOI: 10.1097/ALN.0000000000002694
- [3] Hashimoto DA, Witkowski E, Gao L, Meireles O, Rosman G. Artificial Intelligence in Anesthesiology: Current Techniques, Clinical Applications, and Limitations. *Anesthesiology.* 2020;132(2):379-394. DOI: 10.1097/ALN.0000000000002960
- [4] Paiste, H. J., Godwin, R. C., Smith, A. D., Berkowitz, D. E., & Melvin, R. L. (2024). Strengths-weaknesses-opportunities-threats analysis of artificial intelligence in anesthesiology and perioperative medicine. *Frontiers in Digital Health,* 6. <https://doi.org/10.3389/fdgh.2024.1316931>.
- [5] De Cassai, A., Boscolo, A., Tonetti, T., Ban, I., & Ori, C. (2019). Assignment of ASA-physical status relates to anesthesiologists' experience: a survey-based national-study. *Korean Journal of Anesthesiology,* 72(1), 53–59. <https://doi.org/10.4097/kja.d.18.00224>
- [6] Sankar, A., Johnson, S. R., Beattie, W. S., Tait, G., & Wijeyesundera, D. N. (2014). Reliability of the American Society of Anesthesiologists physical status scale in clinical practice. *British Journal of Anaesthesia,* 113(3), 424–432. <https://doi.org/10.1093/bja/aeu100>.
- [7] Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting Machine Learning predictive models in biomedical research: A multidisciplinary view. *J Med Internet Res.* 2016;18:e323. DOI: 10.2196/jmir.5870
- [8] Georgevici AI, Terblanche M. Neural networks and deep learning: a brief introduction. *Intensive Care Med.* 2019;45:712-714. DOI: 10.1007/s00134-019-05537-w
- [9] Lantz B. *Machine Learning with R: Expert Techniques for Predictive Modeling.* 3rd ed. Packt Publishing; 2019.
- [10] Bellini V, Valente M, Bertorelli G, Pifferi B, Craca M, Mordonini M, et al. Machine learning in perioperative medicine: a systematic review. *J Anesth Analg Crit Care.* 2022;2(1):2. DOI: 10.1186/s44158-022-00033-y
- [11] Bogoń A, Górská M, Ostojńska M, Kałuża I, Dziuba G, Dobosz M. Artificial intelligence in anesthesiology - a review. *J Pre-Clin Clin Res.* 2024;18(3):265-269. DOI: 10.26444/jpccr/191550
- [12] Singhal M, Gupta L, Hirani K. A Comprehensive Analysis and Review of Artificial Intelligence in Anaesthesia. *Cureus.* 2023;15(9):e45038. DOI: 10.7759/cureus.45038
- [13] Shukla A, Salma A, Patel D, John JD, Kantamneni R, Patel T, et al. Harnessing Artificial Intelligence (AI) in Anaesthesiology: Enhancing Patient Outcomes and Clinical Efficiency. *Cureus.* 2024;16(11):e73383. DOI: 10.7759/cureus.73383
- [14] Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* 2019;17(1):195. DOI: 10.1186/s12916-019-1426-2
- [15] Bellini V, Rafano Carnà E, Russo M, Di

- Vincenzo F, Berghenti M, Baciarello M, Bignami E. Artificial intelligence and anesthesia: a narrative review. *Ann Transl Med.* 2022 May;10(9):528. doi: 10.21037/atm-21-7031.
- [16] Srinivasareddy S. Artificial Intelligence in Anesthesia: What Might the Future Hold? *Int J Clin Anesthesiol.* 2024;12(1):1131. DOI: 10.47739/anesthesiology.1131
- [17] Kambale M, Jadhav S. Applications of artificial intelligence in anesthesia: a systematic review. *Saudi J Anaesth.* 2024;18:249-256. 10.4103/sja.sja_955_23
- [18] Adebayo O, Bhuiyan ZA, Ahmed Z. Exploring the effectiveness of artificial intelligence, machine learning and deep learning in trauma triage: A systematic review and meta-analysis. *Digit Health.* 2023;9:20552076231205736. DOI: 10.1177/20552076231205736
- [19] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71. 10.1136/bmj.n71
- [20] Damen JAA, Moons KGM, van Smeden M, et al. How to conduct a systematic review and meta-analysis of prognostic model studies. *Clin Microbiol Infect.* 2023;29:434-440. 10.1016/j.cmi.2022.07.019
- [21] Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med.* 2014;11:e1001744. 10.1371/journal.pmed.1001744
- [22] Ramadanov N, Voss M, Diallo RM, et al. Do Meta-Analyses of Total Hip Arthroplasty Produce Reliable Results? A Systematic Review and Meta-Epidemiological Study of Statistical Methods. *Orthop Surg.* 2025;17:1936-1955. 10.1111/os.70077
- [23] Walker E, Hernandez AV, Kattan MW. Meta-analysis: its strengths and limitations. *Cleve Clin J Med.* 2008;75:431-439. 10.3949/ccjm.75.6.431
- [24] Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med.* 2019;170(1):51-58. DOI: 10.7326/M18-1376
- [25] Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med.* 2019;170(1):W1-W33. DOI: 10.7326/M18-1377
- [26] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366:l4898. DOI: 10.1136/bmj.l4898
- [27] Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:j4008. DOI: 10.1136/bmj.j4008
- [28] Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. *Anaesthesia.* 1984;39:1105-1111. 10.1111/j.1365-2044.1984.tb08932.x
- [29] Savva D. Prediction of difficult tracheal intubation. *Br J Anaesth.* 1994;73:149-153. 10.1093/bja/73.2.149
- [30] Mallampati SR, Gatt SP, Gugino LD. A clinical sign to predict difficult tracheal intubation: a prospective study. *Can Anaesth Soc J.* 1985;32:429-434. 10.1007/bf03011357
- [31] Patil VU, Stehling LC, Zauder HL. *Fiberoptic Endoscopy in Anesthesia.* Chicago: Year Book Medical Publishers; 1983.
- [32] Frerk CM. Predicting difficult intubation. *Anaesthesia.* 1991;46:1005-1008. 10.1111/j.1365-2044.1991.tb09909.x
- [33] Connor CW, Segal S. Accurate classification of difficult intubation by computerized facial analysis. *Anesth Analg.* 2011;112(1):84-93. DOI: 10.1213/ANE.0b013e31820098d6
- [34] Yan HM, Wei XC, Zhang H, Chen XF, Luo EQ. Predicting Cormack classification based on neural network with multiple anthropometric features. *IEEE; 2010.* 10.1109/icacia.2010.5709849
- [35] Koca E, Kutlusoy S, Er MB, Koca T. Predicting Difficult Tracheal Intubation Using Multi-Angle Photographic Analysis with Convolutional Neural Networks and EfficientNet. *Diagnostics.* 2026;16(7):1042. 10.3390/diagnostics16071042
- [36] Tavolara TE, Gurcan MN, Segal S, Niazi MKK. Identification of difficult to intubate patients from frontal face images using an ensemble of deep learning models. *Comput Biol Med.* 2021;136:104737. DOI: 10.1016/j.cbi.2021.104737

10.1016/j.compbimed.2021.104737

[37] 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;9:38. DOI: 10.1186/s40560-021-00551-x

[38] 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. DOI: 10.1186/s12871-021-01343-4

[39] Zhang L, Fabbri D, Lasko TA, Ehrenfeld JM, Wanderer JP. A System for Automated Determination of Perioperative Patient Acuity. *J Med Syst*. 2018;42(7):123. 10.1007/s10916-018-0977-7

[40] Beaulieu-Jones BK, Orzechowski P, Moore JH. Mapping patient trajectories using longitudinal extraction and deep learning in the MIMIC-III critical care database. *bioRxiv*. 2017. 10.1101/177428

[41] Carneiro G, Oakden-Rayner L, Bradley AP, Nascimento J, Palmer L. Automated 5-year mortality prediction using deep learning and radiomics features from chest computed tomography. *ISBI 2017 Apr 1 [cited 2026 Mar];130–4.*; Melbourne. 10.1109/isbi.2017.7950485

[42] Bihorac A, Ozrazgat-Baslanti T, Ebadi A, Motaei A, Madkour M, Pardalos PM, et al. MySurgeryRisk: Development and Validation of a Machine-Learning Risk Algorithm for Major Complications and Death after Surgery. *Ann Surg*. 2019;269(4):652-662. DOI: 10.1097/SLA.0000000000002706

[43] Tseng P, Chen Y, Wang C, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Crit Care*. 2020;24:1-13. 10.1186/s13054-020-03179-9

[44] Xie X, Li J, Zhong Y, et al. A risk prediction model based on machine learning for postoperative cognitive dysfunction in elderly patients with non-cardiac surgery. *Aging Clin Exp Res*. 2023;35(12):2951-2960.

[45] Thottakkara P, Ozrazgat-Baslanti T, Hupf BB, Rashidi P, Pardalos P, Momcilovic P, Bihorac A. Application of Machine Learning Techniques to High-Dimensional Clinical Data to Forecast

Postoperative Complications. *PLoS ONE*. 2016;11(5):e0155705. DOI:

10.1371/journal.pone.0155705

[46] Jalali A, Lonsdale H, Do N, Peck J, Gupta M, Kutty S, et al. Deep Learning for Improved Risk Prediction in Surgical Outcomes. *Sci Rep*. 2020;10:9289. DOI: 10.1038/s41598-020-62971-3

[47] Xue B, Li D, Lu C, et al. Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. *JAMA Netw Open*. 2021;4(3):e212240.

10.1001/jamanetworkopen.2021.2240

[48] Saravi B, Zink A, Ülkümen S, Couillard-Despres S, Hassel F, Lang G. Performance of Artificial Intelligence-Based Algorithms to Predict Prolonged Length of Stay after Lumbar Decompression Surgery. *J Clin Med*. 2022;11:4050. DOI: 10.3390/jcm11144050

[49] Hofer IS, Lee C, Gabel E, Baldi P, Cannesson M. Development and validation of a deep neural network model to predict postoperative mortality, acute kidney injury, and reintubation using a single feature set. *NPJ Digit Med*. 2020;3:58. DOI: 10.1038/s41746-020-0248-0

[50] Fernandes MPB, Armengol de la Hoz M, Rangasamy V, Subramaniam B. Machine Learning Models with Preoperative Risk Factors and Intraoperative Hypotension Parameters Predict Mortality Following Cardiac Surgery. *J Cardiothorac Vasc Anesth*. 2020. DOI: 10.1053/j.jvca.2020.07.029

[51] Hill BL, Brown R, Gabel E, et al. An automated machine learning-based model predicts postoperative mortality using readily-extractable preoperative electronic health record data. *Br J Anaesth*. 2019;123(6):877-886. DOI: 10.1016/j.bja.2019.07.030

[52] Corey KM, Kashyap S, Lorenzi E, Lagoo-Deenadayalan SA, Heller K, Whalen K, et al. Development and validation of machine learning models to identify high-risk surgical patients using automatically curated electronic health record data (Pythia). *PLoS Med*. 2018;15(11):e1002701. DOI: 10.1371/journal.pmed.1002701

[53] Mathis MR, Engoren MC, Joo H, et al. Early Detection of Heart Failure with Reduced Ejection Fraction using Perioperative Data among Non-Cardiac Surgical Patients: A Machine Learning Approach. *Anesth Analg*. 2020;130(5):1188-1200. DOI: 10.1213/ane.0000000000004630

- [54] Flechet M, Falini S, Bonetti C, Güiza F, Schetz M, Van den Berghe G, et al. Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Crit Care*. 2019;23(1):282. DOI: 10.1186/s13054-019-2563-x
- [55] Meiring C, Dixit A, Harris S, et al. Optimal intensive care outcome prediction over time using machine learning. *PLoS One*. 2018;13(11):e0206862. 10.1371/journal.pone.0206862
- [56] Pang K, Li L, Ouyang W, Liu X, Tang Y. Establishment of ICU mortality risk prediction models with machine learning algorithm using MIMIC-IV database. *Diagnostics*. 2022;12:1068. 10.3390/diagnostics12051068
- [57] Safaei N, Safaei B, Seyedekrami S, et al. E-CatBoost: An efficient machine learning framework for predicting ICU mortality using the eICU Collaborative Research Database. *PLoS One*. 2022;17:e0262895. 10.1371/journal.pone.0262895
- [58] Lei Y, Zhang Z, Li Y, et al. Machine learning-based prognostic model for in-hospital mortality of aortic dissection: insights from an intensive care medicine perspective. *Digit Health*. 2024 Jan 1 [cited 2026 May];10; 10.1177/20552076241269450
- [59] Sun Y, He Z, Ren J, Wu Y. Prediction model of in-hospital mortality in intensive care unit patients with cardiac arrest: a retrospective analysis of MIMIC-IV database based on machine learning. *BMC Anesthesiol*. 2023;23:178. 10.1186/s12871-023-02138-5
- [60] Dhami A, Onyeukwu KA, Sattar S, Batra A, Mostafa Y, Haris M, et al. The Prognostic Performance of Artificial Intelligence and Machine Learning Models for Mortality Prediction in Intensive Care Units: A Systematic Review. *Cureus*. 2025;17(8):e90465. 10.7759/cureus.90465
- [61] Hemmerling TM. Robots will perform anesthesia in the near future. *Anesthesiology*. 2020;132:219-220. 10.1097/aln.0000000000003435
- [62] Parvinian B, Scully C, Wiyor H, Kumar A, Weininger S. Regulatory considerations for physiological closed-loop controlled medical devices used for automated critical care. *Anesth Analg*. 2018;126:1916-1925. 10.1213/ANE.0000000000002329
- [63] Pasin L, Nardelli P, Pintaudi M, Greco M, Zambon M, Cabrini L, Zangrillo A. Closed-Loop Delivery Systems Versus Manually Controlled Administration of Total IV Anesthesia: A Meta-analysis of Randomized Clinical Trials. *Anesth Analg*. 2017;124(2):456-464. 10.1213/ane.0000000000001394
- [64] Hemmerling TM, Charabati S, Zaouter C, Minardi C, Mathieu PA. A randomized controlled trial demonstrates that a novel closed-loop propofol system performs better hypnosis control than manual administration. *Can J Anaesth*. 2010;57:725-35. 10.1007/s12630-010-9335-z
- [65] Hemmerling TM, Arbeid E, Wehbe M, Cyr S, Taddei R, Zaouter C. Evaluation of a novel closed-loop total intravenous anaesthesia drug delivery system: a randomized controlled trial. *Br J Anaesth*. 2013;110(6):1031-1039. 10.1093/bja/aet001
- [66] Dussaussoy C, Peres M, Jaoul V, Liu N, Chazot T, Picquet J, et al. Automated titration of propofol and remifentanyl decreases the anesthesiologist's workload during vascular or thoracic surgery: a randomized prospective study. *J Clin Monit Comput*. 2013 Mar 14 [cited 2026 May];28(1):35–40. 10.1007/s10877-013-9453-6.
- [67] De Smet T, Struys MM, Neckebroek MM, Van den Hauwe K, Bonte S, Mortier EP. The Accuracy and Clinical Feasibility of a New Bayesian-Based Closed-Loop Control System for Propofol Administration Using the Bispectral Index. *Anesth Analg*. 2008 Sept 16 [cited 2026 Apr];107(4):1200–10. 10.1213/ane.0b013e31817bd1a6
- [68] Liu N, Chazot T, Genty A, et al. Titration of Propofol for Anesthetic Induction and Maintenance Guided by the Bispectral Index: Closed-loop versus Manual Control: A Prospective, Randomized, Multicenter Study. *Anesthesiology*. 2006;104(4):686-695. 10.1097/00000542-200604000-00012
- [69] Liu N, Chazot T, Hamada S, Landais A, Boichut N, Dussaussoy C, et al. Closed-Loop Coadministration of Propofol and Remifentanyl Guided by Bispectral Index: A Randomized Multicenter Study. *Anesth Analg*. 2011;112(3):546-557. 10.1213/ane.0b013e318205680b
- [70] Liu Y, Li M, Yang D, et al. Closed-Loop Control Better than Open-Loop Control of Propofol TCI Guided by BIS: A Randomized, Controlled, Multicenter Clinical Trial to Evaluate the CONCERT-CL Closed-Loop System. *PLoS One*. 2015;10(4):e0123862. 10.1371/journal.pone.0123862
- [71] Puri GD, Kumar B, Aveek J. Closed-loop anaesthesia delivery system (CLADS) using bispectral index: a performance assessment study.

Anaesth Intensive Care. 2007;35:357-362. 10.1177/0310057x0703500306

[72] Puri GD, Mathew PJ, Biswas I, et al. A Multicenter Evaluation of a Closed-Loop Anesthesia Delivery System: A Randomized Controlled Trial. *Anesth Analg*. 2016;122(1):106-114. 10.1213/ANE.0000000000000769

[73] Struys MM, De Smet T, Versichelen LF, Van de Velde S, Van den Broecke R, Mortier EP. Comparison of Closed-loop Controlled Administration of Propofol Using Bispectral Index as the Controlled Variable versus "Standard Practice". *Anesthesiology*. 2001;95:6-17. 10.1097/00000542-200107000-00007

[74] Morley A, Derrick J, Mainland P, Lee BB, Short TG. Closed loop control of anaesthesia: an assessment of the bispectral index as the target of control. *Anaesthesia*. 2000;55(10):953-959. 10.1046/j.1365-2044.2000.01527.x

[75] Ngan Kee WD, Tam YH, Khaw KS, Ng FF, Lee SWY. Closed-Loop Feedback Computer-Controlled Phenylephrine for Maintenance of Blood Pressure During Spinal Anesthesia for Cesarean Delivery. *Anesth Analg*. 2017;125(1):117-123. 10.1213/ane.0000000000001974

[76] Xu C, Zhu Y, Wu L, et al. Evaluating the effect of an artificial intelligence system on the anesthesia quality control during gastrointestinal endoscopy with sedation: a randomized controlled trial. *BMC Anesthesiol*. 2022;22(1):1796. 10.1186/s12871-022-01796-1

[77] Zaouter C, Taddei R, Wehbe M, Arbeid E, Cyr S, Giunta F, Hemmerling TM. A novel system for automated propofol sedation: hybrid sedation system (HSS). *J Clin Monit Comput*. 2016 Mar 12 [cited 2026 May];31(2):309–17. 10.1007/s10877-016-9858-0.

[78] Joosten A, Jame V, Alexander B, et al. Feasibility of Fully Automated Hypnosis, Analgesia, and Fluid Management Using 2 Independent Closed-Loop Systems During Major Vascular Surgery: A Pilot Study. *Anesth Analg*. 2018 May 10 [cited 2026 Feb];128(6). 10.1213/ane.0000000000003433.

[79] Absalom AR, Sutcliffe N, Kenny GN. Closed-loop Control of Anesthesia Using Bispectral Index: Performance Assessment in Patients Undergoing Major Orthopedic Surgery. *Anesthesiology*. 2002;96(1):67-73. 10.1097/00000542-200201000-00017

[80] Hemmerling TM, Arbeid E, Wehbe M, Cyr S,

Giunta F, Zaouter C. Transcontinental anaesthesia: a pilot study. *Br J Anaesth*. 2013;110(5):758-763. 10.1093/bja/aes498

[81] Ren W, Chen J, Liu J, et al. Feasibility of intelligent drug control in the maintenance phase of general anesthesia based on convolutional neural network. *Heliyon*. 2023;9(1):e12481. [10.1016/j.heliyon.2022.e12481](https://doi.org/10.1016/j.heliyon.2022.e12481)

[82] Wehbe M, Arbeid E, Cyr S, Mathieu PA, Taddei R, Morse J, Hemmerling TM. A technical description of a novel pharmacological anesthesia robot. *J Clin Monit Comput*. 2013 June 14 [cited 2026 Jan];28(1):27–34. 10.1007/s10877-013-9451-8.

[83] 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*. 2021 Dec 2 [cited 2026 Apr];123:102227–102227. Available from: <https://doi.org/10.1016/j.artmed.2021.102227>

[84] Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, et al. Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. *Anesthesiology*. 2018;129:663-674. 10.1097/aln.0000000000002300

[85] Kendale S, Kulkarni P, Rosenberg AD, Wang J. Supervised Machine Learning Predictive Analytics for Prediction of Postinduction Hypotension. *Anesthesiology*. 2018;129:675-688. 10.1097/aln.0000000000002374

[86] 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 Apr 14 [cited 2026 Apr];130(5):1157–66. 10.1213/ane.0000000000004539.

[87] 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 Apr 14 [cited 2026 Apr];130(5):1201–10. 10.1213/ane.0000000000004636.

[88] Miyaguchi N, Takeuchi K, Kashima H, Morita M, Morimatsu H. Predicting anesthetic infusion events using machine learning. *Sci Rep*. 2021;11(1):23648. 10.1038/s41598-021-03112-2

[89] Bowness J, Varsou O, Turbitt L, Burckett-St Laurent D. Identifying anatomical structures on

- ultrasound: assistive artificial intelligence in ultrasound-guided regional anesthesia. *Clin Anat.* 2021;34(5):802-809. 10.1002/ca.23742
- [90] Bowness JS, Macfarlane AJR, Burckett-St Laurent D, et al. Evaluation of the impact of assistive artificial intelligence on ultrasound scanning for regional anaesthesia. *Br J Anaesth.* 2023;130(2):226-233. DOI: 10.1016/j.bja.2022.07.049
- [91] Bowness JS, El-Boghdady K, Woodworth G, et al. Exploring the utility of assistive artificial intelligence for ultrasound scanning in regional anesthesia. *Reg Anesth Pain Med.* 2022;47:375-379. DOI: 10.1136/rapm-2021-103368
- [92] McKendrick M, Yang S, McLeod GA. The use of artificial intelligence and robotics in regional anaesthesia. *Anaesthesia.* 2021;76(Suppl 1):171-181. 10.1111/anae.15274
- [93] Cai N, Wang G, Xu L, Zhou Y, Chong H, Zhao Y, et al. Examining the impact of perceptual learning artificial-intelligence-based on the incidence of paresthesia when performing the ultrasound-guided popliteal sciatic block. *BMC Anesthesiol.* 2022;22:392. DOI: 10.1186/s12871-022-01937-6
- [94] Kowa CY, Morecroft M, Macfarlane AJR, et al. Prospective randomized evaluation of the sustained impact of assistive artificial intelligence on anesthetists' ultrasound scanning for regional anesthesia. *BMJ Surg Interv Health Technologies.* 2024 Oct 1 [cited 2026 May];6(1). 10.1136/bmjst-2024-000264.
- [95] Gungor I, Gunaydin B, Oktar SO, et al. A real-time anatomy identification via tool based on artificial intelligence for ultrasound-guided peripheral nerve block procedures: an accuracy study. *J Anesth.* 2021;35(4):591-594. 10.1007/s00540-021-02947-3.
- [96] Hetherington J, Lessoway V, Gunka V, Abolmaesumi P, Rohling R. SLIDE: automatic spine level identification system using a deep convolutional neural network. *Int J Comput Assist Radiol Surg.* 2017 Mar 30 [cited 2026 Apr];12(7):1189–98. DOI: 10.1007/s11548-017-1575-8
- [97] Leng Y, Yu S, Tan KK, Tildsley P, Sia AT, Sng BL. Development of a Real-time Lumbar Ultrasound Image Processing System for Epidural Needle Entry Site Localization. *IEEE EMBC.* 2016 Aug 1 [cited 2025 Nov];2016:4093–6. 10.1109/embc.2016.7591626.
- [98] Ikhsan M, Tan KK, Oh TT, Lew JP, Sng BL. Gabor-based Automatic Spinal Level Identification in Ultrasound. *IEEE.* 2017 July 1 [cited 2025 Nov];2017:3146–9. 10.1109/embc.2017.8037524.
- [99] Pesteie M, Lessoway V, Abolmaesumi P, Rohling RN. Automatic localization of the needle target for ultrasound-guided epidural injections. *IEEE Trans Med Imaging.* 2018 Aug 11 [cited 2025 Nov];37(1):81–92. 10.1109/tmi.2017.2739110.
- [100] 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: 116–116. 10.1186/s12871-021-01331-8
- [101] Morse J, Terrasini N, Wehbe M, et al. Comparison of success rates, learning curves, and inter-subject performance variability of robot-assisted and manual ultrasound-guided nerve block needle guidance in simulation. *Br J Anaesth.* 2014;112(6):1092-1097. 10.1093/bja/aet440
- [102] Hayashida, M., Sato, M., Nagashima, M., Hasegawa, M., Katori, N., & Yin, O. Q. (2015). Prediction formulas for individual opioid analgesic requirements based on genetic polymorphism analyses. *PLOS ONE*, 10(1), e0116885. <https://doi.org/10.1371/journal.pone.0116885>
- [103] Nair AA, Velagapudi MA, Lang JA, Behara L, Venigandla R, Velagapudi N, et al. Machine learning approach to predict postoperative opioid requirements in ambulatory surgery patients. *PLoS ONE.* 2020;15(7):e0236833. DOI: 10.1371/journal.pone.0236833
- [104] Dolendo IM, Wallace AM, Armani A, Waterman RS, Said ET, Gabriel RA. Predictive Analytics for Inpatient Postoperative Opioid Use in Patients Undergoing Mastectomy. *Cureus.* 2022. DOI: 10.7759/cureus.23079
- [105] Gabriel RA, Harjai B, Prasad RS, Simpson S, Chu I, Fisch KM, Said ET. Machine learning approach to predicting persistent opioid use following lower extremity joint arthroplasty. *Reg Anesth Pain Med.* 2022 May;47(5):313-319. doi: 10.1136/rapm-2021-103299
- [106] Liu R, Gutiérrez R, Mather RV, Stone TAD, Santa Cruz Mercado LA, et al. Development and prospective validation of postoperative pain prediction from preoperative EHR data using attention-based set embeddings. *NPJ Digit Med.* 2023. DOI: 10.1038/s41746-023-00947-z
- [107] Hu YJ, Ku TH, Jan RH, Wang K, Tseng YC, Yang SF. Decision tree-based learning to predict

patient controlled analgesia consumption and readjustment. *BMC Med Inform Decis Mak.* 2012;12:131. 10.1186/1472-6947-12-131

[108] Shim JG, Ryu KH, Cho EA, Ahn JH, Cha YB, Lim G, et al. Machine learning for prediction of postoperative nausea and vomiting in patients with intravenous patient-controlled analgesia. *PLoS ONE.* 2022;17(12):e0277957. DOI: 10.1371/journal.pone.0277957

[109] 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): 115–115. 10.1186/s12871-020-01036-4

[110] Kartal E, Koçoglu FO, Özen Z, Emre İE, Güngör G, Bozkuş PS. An Intelligent Postoperative Chronic Pain Prediction System (I-POCPP). *J Ist Faculty Med.* 2022 June 13 [cited 2026 May];85(3):416–24. 10.26650/iuitfd.972738.

[111] Sun C, Li M, Lan L, et al. Prediction models for chronic postsurgical pain in patients with breast cancer based on machine learning approaches. *Front Oncol.* 2023;13:1096468. 10.3389/fonc.2023.1096468

[112] Aladro Larenas XM, Castillo Cuadros M, Miguel Aranda IE, et al. Postoperative Pain at Discharge From the Post-anesthesia Care Unit: A Case-Control Study. *Cureus.* 2024 Oct 24 [cited 2026 Apr];16(10). 10.7759/cureus.72297.

[113] Zhang C, He J, Liang X, et al. Deep learning models for the prediction of acute postoperative pain in PACU for video-assisted thoracoscopic surgery. *BMC Med Res Methodol.* 2024;24:232. 10.1186/s12874-024-02357-5

[114] Shi G, Liu G, Gao Q, et al. A random forest algorithm-based prediction model for moderate to severe acute postoperative pain after orthopedic surgery under general anesthesia. *BMC Anesthesiol.* 2023;23:361. 10.1186/s12871-023-02328-1

[115] Fuica R, Krochek C, Weissbrod R, Greenman D, Freundlich A, Gozal Y. Reduced postoperative pain in patients receiving nociception monitor guided analgesia during elective major abdominal surgery: a randomized, controlled trial. *J Clin Monit Comput.* 2023;37:481-491. DOI: 10.1007/s10877-022-00906-1

[116] 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):1070-1078. DOI: 10.1016/j.bja.2020.07.057

[117] Park I, Park JH, Koo BW. Machine learning model of facial expression outperforms models using analgesia nociception index and vital signs to predict postoperative pain intensity: a pilot study. *Korean J Anesthesiol.* 2024;77(2):195-204. 10.4097/kja.23583

[118] Wu CL, Liu SF, Yu TL, Shih SJ, Chang CH, et al. Deep Learning-Based Pain Classifier Based on the Facial Expression in Critically Ill Patients. *Front Med.* 2022;9:851690. DOI: 10.3389/fmed.2022.851690

[119] Bargshady G, Zhou X, Deo RC, Soar J, Whittaker F, Wang H. Enhanced deep learning algorithm development to detect pain intensity from facial expression images. *Expert Syst Appl.* 2020 Feb 16 [cited 2026 May];149:113305–113305. 10.1016/j.eswa.2020.113305.

[120] 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. 10.2196/23920

[121] Susam BT, Akcakaya M, Nezamfar H, et al. Automated Pain Assessment using Electrodermal Activity Data and Machine Learning. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:372-375. 10.1109/embc.2018.8512389

[122] Bellal M, Lelandais J, Chabin T, et al. Calibration trial of an innovative medical device (NEVVAD) for the evaluation of pain in non-communicating patients in the intensive care unit. *Front Med.* 2024;11:1309720. 10.3389/fmed.2024.1309720

[123] 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. 10.1371/journal.pone.0238249

[124] 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):3408-3415. 10.1109/jbhi.2021.3068481

[125] Shalbaf A, Saffar M, Sleight JW, Shalbaf R. Monitoring the depth of anesthesia using a new adaptive neuro-fuzzy system. *IEEE J Biomed*

Health Inform. 2018 May;22(3):671-677. doi: 10.1109/JBHI.2017.2709841.

[126] Gu Y, Liang Z, Hagihira S. Use of Multiple EEG Features and Artificial Neural Network to Monitor the Depth of Anesthesia. *Sensors*. 2019;19(11):2499. 10.3390/s19112499

[127] Belur Nagaraj S, Ramaswamy SM, Weerink MAS, Struys MMRF. Predicting Deep Hypnotic State From Sleep Brain Rhythms Using Deep Learning: A Data-Repurposing Approach. *Anesth Analg*. 2020;130(5):1211-1221. 10.1213/ANE.0000000000004651

[128] Nair BG, Newman SF, Peterson GN, Schwid HA. Smart Anesthesia Manager (SAM)-A Real-time Decision Support System for Anesthesia Care during Surgery. *IEEE Trans Biomed Eng*. 2013 Jan;60(1):207-10. doi: 10.1109/TBME.2012.2205384.

[129] Rinehart J, Chung E, Canales C, Cannesson M. Intraoperative Stroke Volume Optimization Using Stroke Volume, Arterial Pressure, and Heart Rate: Closed-Loop (Learning Intravenous Resuscitator) Versus Anesthesiologists. *J Cardiothorac Vasc Anesth*. 2012;26(6):933-939. 10.1053/j.jvca.2012.05.015

[130] Canales C, Lee C, Cannesson M. Science Without Conscience Is but the Ruin of the Soul: The Ethics of Big Data and Artificial Intelligence in Perioperative Medicine. *Anesth Analg*. 2020;130(5):1234-1243. DOI: 10.1213/ANE.0000000000004728

[131] Char DS, Burgart A. Machine Learning Implementation in Clinical Anesthesia: Opportunities and Challenges. *Anesth Analg*. 2020;130(6):1709-1712. DOI: 10.1213/ANE.0000000000004656

[132] Mökander J. Auditing of AI: Legal, ethical and technical approaches. *Digit Soc*. 2023;2(3):49. 10.1007/s44206-023-00074-y

[133] Joshi S, Sharma M, Das RP, et al. Modeling Conceptual Framework for Implementing Barriers of AI in Public Healthcare. *Sustainability*. 2022;14:11698. 10.3390/su141811698

[134] Reddy S. Generative AI in healthcare: An implementation science informed translational path on application, integration and governance. *Implement Sci*. 2024;19(1):27. 10.1186/s13012-024-01357-9

[135] Ueda D, Kakinuma T, Fujita S, et al. Fairness of artificial intelligence in healthcare: Review and recommendations. *Jpn J Radiol*. 2024;42(1):3-15.

10.1007/s11604-023-01474-3

[136] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177-188. DOI: 10.1016/0197-2456(86)90046-2