



Artificial Intelligence and Machine Learning in Pediatric Anesthesia: A Narrative Review

Patrick Fakhoury MS4¹, Aditya Shah MS3¹, Caleb Zimmerman BS¹, Aiman Almasnaah MS4¹, Kush Patel MS4¹, Deepti Sanku MS4¹, Kabir Mendiratta MD⁴, Wael Saasouh MD^{2,3}

¹Central Michigan University College of Medicine, Saginaw, MI, United States

²Wayne State University School of Medicine, Detroit, MI, USA

³Outcomes Research Consortium, Houston, TX, USA

⁴Detroit Medical Center, Detroit, MI, USA



Introduction

- Pediatric anesthesia presents unique challenges: small airway dimensions, rapid desaturation, limited physiologic reserve, and difficulty with postoperative pain assessment.
- Traditional approaches such as formula-based ETT sizing and clinician judgment often lack precision and consistency.
- Artificial intelligence (AI) and machine learning (ML) are increasingly being investigated to support perioperative decision-making in children.
- Applications span multiple domains including airway management, intraoperative risk prediction, and postoperative monitoring.

Objective

- To review and synthesize current applications of AI and ML in pediatric anesthesia and evaluate their potential role in perioperative care.

Methods

Design: Narrative review following PRISMA and search strategy

Databases searched: EBSCO CINAHL Plus, PubMed, Scopus, Web of Science

Search terms: “pediatric anesthesia,” “artificial intelligence,” “machine learning,” “airway,” “endotracheal tube,” “hypoxemia,” “pain assessment,” “perioperative monitoring,” etc.

Inclusion criteria:

- Studies applying AI/ML in pediatric anesthesia (0–18 years)
- Outcomes related to airway management, intraoperative monitoring, or postoperative care
- Original research, English language

Exclusion criteria:

- Adult-only or mixed-age studies without pediatric subgroup data
- Simulation-only or purely theoretical modeling without clinical dataset

Data extraction: Structured table capturing population, setting, sample size, AI/ML method, comparator (if any), target outcome, and performance metrics.

Final Yield: Records Identified 893 papers -> Screened -> included 11 papers confirmed by 2 separate reviewers.

Results

Study	Year	Country	Design	Sample Size	Application Domain	Conventional Methods	AI/ML	Data Modality
Jalali et al.	2016	USA	Algorithm design	600	Automatic detection of intubation time, early malposition detection	Manual Monitoring	Signal-processing	EtCO ₂ , PIP, TV (q15s)
Galvez et al.	2017	USA	Retrospective observational	900	Automatic Airway device classification (mask, LMA, ETT)	Manual device classification	NN, SVM, Boosted Trees	EtCO ₂ , TV, PIP, RR; minute ventilation (@ 1 min)
Zhou et al.	2022	China	Retrospective, multicenter	990	ML prediction of ETT size (cuffed vs uncuffed)	Age/weight-based ETT formulas	RF, SVR, GBR	Age, height, weight, BMI, CXR tracheal diameters + carina length
Im et al.	2022	South Korea	Retrospective cohort	630	Prediction of BiPAP failure requiring intubation	Standard bedside monitoring	LSTM-RNN	Vitals, labs, meds/pressors/sedatives, respiratory support & settings
Kim et al.	2023	South Korea	Retrospective cohort	37,057	ETT size/depth prediction (cuffed vs uncuffed)	Age-based ETT formulas	GBRT, MLR	Age, sex, height, and weight
Park et al.	2023	South Korea	Retrospective observational	13,130	Hypoxemia prediction	Clinician judgement, vital sign thresholds	GBM, LSTM, Transformer	SpO ₂ , EtCO ₂ , FiO ₂ , TV, PIP, age, sex, height, and weight
Gude et al.	2024	Norway	Retrospective observational	149	Postoperative pain assessment with indication for rescue medication	Pain Scores (CHIPPS, PPPM-D scales)	Symbolic regression, Boruta, CNN	CHIPPS, PPPM-D scales, and time after surgery
Shim et al.	2023	South Korea	Retrospective cohort	859	ETT placement prediction	Age-based ETT formula	RF, SVM, ANN, Elastic Net	Age, sex, height, weight
Shim et al.	2021	South Korea	Retrospective cohort	834	ETT depth optimization	Age-based ETT formula	RF, EN, SVM, ANN	Age, sex, height, weight
Grey et al.	2023	European countries	Prospective cohort	30,325	ML decision support & risk stratification based on APRICOT	ASA-PS	GBM	EHR variables + Procedure time, procedure typ, patient type, anesthesia type
Yu et al.	2023	USA	Retrospective cohort	54,776	Predicting pediatric emergence delirium using ML on EHR	No single bedside risk score	LASSO, ridge, RF, XGB	EHR variables from pre-op, intra-op, and PACU

Table 1: Characteristics of included studies (n = 11) on AI/ML applications in pediatric anesthesia.

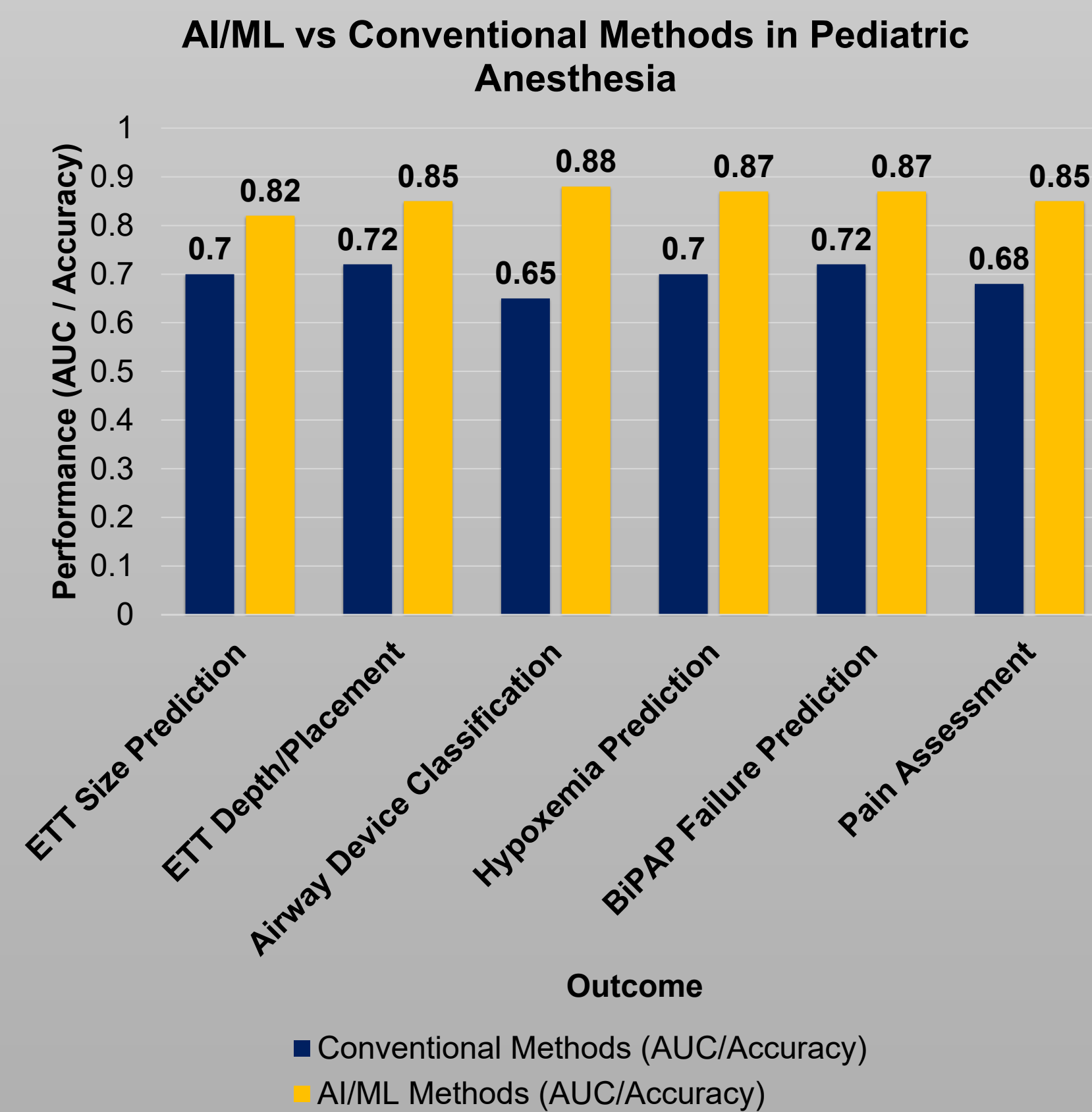


Figure 1. Comparison of predictive performance (AUC/Accuracy) between conventional methods and AI/ML across six pediatric anesthesia applications. An AUC of 0.5 indicates that the model performs no better than random chance. Values between **0.7** and **0.8** represent **acceptable discrimination** or modest predictive power. AUC values from **0.8** to **0.9** indicate **good discrimination**, while values ≥ 0.9 reflect **excellent discrimination**, which is rarely achieved in clinical models.

Paper	One-sentence finding
Jalali et al., 2016	Wavelet-based EtCO ₂ features plus PIP/TV jumps detected pediatric intubation timing enabling accurate real-time event timestamps
Gálvez et al., 2017	A multilayer neural net on minute-wise ventilator signals distinguished mask vs invasive (LMA/ETT)
Zhou et al., 2022	Random forest using age/anthropometrics plus CXR tracheal measurements predicted cuffed/uncuffed ETT size better than age formulas
Im et al., 2022	A time-aware LSTM on PICU EMR predicted BiPAP failure at 6 h/24 h, outperforming logistic models and enabling earlier escalation.
Kim et al., 2023	On 37k+ cases, GBRT predicted ETT size and depth better than age rules and linear regression, with a simple, deployable feature set.
Park et al., 2023	High-frequency monitor/ventilator signals with gradient boosting predicted SpO₂ <95% one minute ahead supporting earlier intraoperative intervention.
Gude et al., 2024	Minimal 4-predicate rule reproduces the rescue-medication criterion with ~95% CV accuracy, stays informative when CHIPPS alone fades, and is faster to use than full PPPM-D/CHIPPS.
Shim et al., 2023	Elastic-net from demographics cut inappropriate depth to ~17.9% (vs 35.7–62.2% with formulas), markedly lowering deep/endorchinal placements in <7-year-olds.
Shim et al., 2021	Demographics-only ML (RF) placed more tubes in the T1–T3 ‘ optimal zone ’ (~79%) than age/height/ID formulas (44–67%), reducing shallow/deep placements.
Gray et al., 2023	Gradient-boosting models identified low-risk children with NPV >95% at booking and day-of-surgery (AUROC ~0.62/0.72) using APRICOT variables, suggesting triage utility.
Yu et al., 2023	EHR-based ML (LASSO/RF/XGB) predicted pediatric emergence delirium with AUROC ~0.74–0.75 and highlighted modifiable perioperative factors such as premedication (midazolam), antiemetics (ondansetron), induction techniques (IV induction, lower risk), analgesia (ketorolac – lower risk), and state on PACU arrival (asleep/unconscious).

Table 2. Key findings of 11 pediatric perioperative AI/ML studies spanning airway management (ETT size/depth/placement), intraoperative monitoring (hypoxemia), postoperative outcomes (pain, emergence delirium), and perioperative risk stratification.

Discussion

Key Insights:

- AI/ML demonstrated **superior performance** compared to conventional methods for ETT sizing/placement, hypoxemia prediction, BiPAP failure, and pain assessment.
- The greatest gains were observed in **airway device classification** and **hypoxemia prediction**, where formula-based or threshold-based methods were less reliable.
- Importantly, AI/ML tools are best viewed as **clinical decision-support systems**, augmenting anesthesiologists’ expertise, not replacing physician judgment.

Clinical Relevance:

- Improved prediction of airway size and depth **reduces complications** such as hypoxemia and airway trauma.
- Early recognition** of desaturation risk through AI/ML allows timely intervention, potentially improving safety in high-risk pediatric cases.
- Automated pain recognition and BiPAP failure prediction provide **objective monitoring**, especially valuable in children with limited ability to communicate discomfort

Limitations:

- Majority of studies were retrospective, single-center, and concentrated in Asia, limiting generalizability.
- Small datasets in certain applications (e.g., pain assessment, BiPAP prediction) reduce robustness.
- Lack of prospective, multicenter validation and heterogeneity in outcome definitions remain barriers to clinical translation.

Conclusion

Artificial intelligence and machine learning show strong potential to enhance perioperative safety and care quality in pediatric anesthesia. Across applications, AI/ML models have outperformed conventional methods, identified key predictors and improved outcome prediction. These technologies should be understood as complementary tools that support, not replace, physicians to offer more accurate information to guide airway management, intraoperative monitoring, and postoperative care. With prospective validation and integration into clinical workflows, AI/ML has the potential to become a trusted tool in optimizing anesthesia care for children.

PRISMA



References

