*Review*

# Artificial intelligence in anesthesiology

F. Gheysen[1], S. Rex[2]

*[1]University Hospitals of Leuven, Department of Anesthesiology, Herestraat 49, 3000 Leuven, Belgium; [2]University Hospitals of Leuven, Department of Anesthesiology, Herestraat 49, 3000 Leuven, Belgium. Associate Professor of Anesthesia, Dpt. of Cardiovascular Sciences, KU Leuven.*

Corresponding author: Gheysen Flore, UZ Leuven, Herestraat 49, 3000 Leuven, Belgium. E-mail: flore.gheysen@uzleuven.be

## *Abstract*

**Artificial intelligence (AI) is rapidly evolving and gaining attention in the medical world. Our aim is to provide readers with insights into this quickly changing medical landscape and the role of clinicians in the middle of this popular technology. In this review, our aim is to explain some of the increasingly frequently used AI terminology explicitly for physicians. Next, we give a summation, an overview of currently existing applications, future possibilities for AI in the medical field of anesthesiology and thoroughly highlight possible problems that could arise from implementing this technology in daily practice.**
**Therefore, we conducted a literature search, including all types of articles published between the first of January 2010 and the 1st of May 2023, written in English, and having a free full text available. We searched Pubmed, Medline, and Embase using "artificial intelligence", "machine learning", "deep learning", "neural networks" and "anesthesiology" as MESH terms.**
**To structure these findings, we divided the results into five categories: preoperatively, perioperatively, postoperatively, AI in the intensive care unit and finally, AI used for teaching purposes. In the first category, we found AI applications for airway assessment, risk prediction, and logistic support. Secondly, we made a summation of AI applications used during the operation. AI can predict hypotensive events, delivering automated anesthesia, reducing false alarms, and aiding in the analysis of ultrasound anatomy in locoregional anesthesia and echocardiography. Thirdly, namely postoperatively, AI can be applied in predicting acute kidney injury, pulmonary complications, postoperative cognitive dysfunction and can help to diagnose postoperative pain in children.**
**At the intensive care unit, AI tools discriminate acute respiratory distress syndrome (ARDS) from pulmonary oedema in pleural ultrasound, predict mortality and sepsis more accurately, and predict survival rates in severe Coronavirus-19 (COVID-19). Finally, AI has been described in training residents in spinal ultrasound, simulation, and plexus block anatomy.**
**Several concerns must be addressed regarding the use of AI. Firstly, this software does not explain its decision process (i.e., the 'black box problem'). Secondly, to develop AI models and decision support systems, we need big and accurate datasets, unfortunately with potential unknown bias. Thirdly, we need an ethical and legal framework before implementing this technology. At the end of this paper, we discuss whether this technology will be able to replace the clinician one day.**
**This paper adds value to already existing literature because it not only offers a summation of existing literature on AI applications in anesthesiology but also gives clear definitions of AI itself and critically assesses implementation of this technology.**

*Keywords:* **Artificial Intelligence, Anesthesia, Machine Learning, Deep Learning, Neural networks, Chatbots, Automation, Decision support systems**

*Ethics committee : Approval of the Ethical Research Committee UZ/KU Leuven was obtained on 04/11/2021.*

## Introduction

### *Why AI?*

The famous AI chatbot "ChatGPT" (OpenAI, San Francisco, United States) successfully passed the threshold of 60% at the United States Medical Licensing Exam (USMLE) without input from human trainers[1]. Furthermore, responses of this chatbot to anonymous medical questions on public social media forums were evaluated by a jury and compared to answers from physicians. Chatbot responses were classified as scientifically very good and more empathetic, now outperforming clinicians[2]. Since anesthesiology heavily relies on technology, AI found its way quickly to this branch of healthcare, where automation has already been proved to improve patient care[3].

This paper aims to give a summation, an overview of currently existing applications for AI in the medical field of anesthesiology and thoroughly dig deeper into some problems that arise from implementing this technology in daily practice. Our goal is to provide readers with insights into this quickly changing medical landscape and the future role of clinicians in the middle of popular technology.

The number of machine learning (ML) clinical trials has gradually increased in the past decade[4], but the idea of automation is not new at all. Bickford introduced the first automated anesthesia system using EEG signals in 1950[5]. It automatically delivered barbiturate and ether anesthesia in a rabbit, cat, monkey and man[6]. Over half a century later, McSleepy, an invention of the McGill University in Montréal, Quebec, Canada, was introduced with a similar feedback control system[6]. AI, however, is much more than solely a closed-loop feedback mechanism[6], which we will explain further on.

### *Definitions*

We summed up several definitions to gain a clear understanding of the basic concepts of AI. Terms such as 'artificial intelligence' and 'machine learning' or 'deep learning' are often used interchangeably, frequently leading to confusion[7].
Artificial Intelligence is a subfield within computer science associated with constructing machines and computers capable of intelligent behavior, simulating the human decision capacity[7].

Machine learning is a part of AI, a subset. It can learn from experience without being programmed explicitly. All ML models are AI models by default, but not vice versa. ML is divided into supervised vs. unsupervised learning, hybrid semi-supervised learning and reinforcement learning[8]. Supervised learning trains the model with the correct answer, meaning the expert clinician gives the label, for example, annotated images. In unsupervised learning, the machine will reorganize data in clusters of similarities[9]. In reinforcement learning, models train with a mathematical reward system, learn from their mistakes, and adapt.

Deep learning (DL) is a subdivision of machine learning that recognizes patterns in data. This technique allows results obtained with neural networks in various fields. Artificial neural network architectures rely on neural layers, imitating the human brain[6]. In contrast to shallow neural networks, deep neural networks consist of more than two layers and are further divided by their architecture into recurrent neural networks (RNN)[7], convolutional neural networks (CNN), and many more. Unlike other types of machine learning, deep learning does not require humans to interfere with the training process. Although the most popular ML method, deep learning is not the only machine learning model. Other examples are linear regression, logistic regression, decision trees, naïve Bayes, support vector machines, linear discriminant analysis etc. These techniques are sometimes called classical machine-learning models. ChatGPT (Open AI, San Francisco, United States) is based on a transformer architecture within AI. This technique excels in applying context in a sentence and in parallelization, which causes its speed.

How well machine learning models perform is often indicated by AUC (AUROC, area under the receiver operating characteristics). Precision determines the percentage of correct positive predictions (true positive rate). Accuracy describes more generally how the model performs across all classes. It is calculated by dividing the number of correct predictions by the total number of predictions.

## Methods

We searched Pubmed, Medline and Embase using MESH-terms "artificial intelligence", "deep learning", "neural networks" OR "machine learning" AND "anesthesia", "anesthesiology", "intensive care" in the title or abstract. We included all types of articles. Eligible articles were published between the first of January 2010 until the 1st of May 2023, written in English, and had a free full text available.

## Results

### *Included articles*

We found 73 articles, of which 8 were excluded based on their relevance. We included one less recent article because of its historical relevance. Later, we added a handful of articles, published between February and April of 2023.

### *Preoperative use of AI*

Before surgery, it is essential to assess possible risks and complications and yet maintain efficiency. By this, admission to the intensive care unit, delayed recovery, and prolonged hospital stay can be predicted. AI can help to better assess airway risks and make predictions more accurate than current risk assessments, thereby making healthcare safer. However, it is also essential that the workflow in the operating theatre remains efficient and that productivity is guaranteed within this safety framework. AI can also aid at this logistic level.

#### *Risk stratification*

AI can help with risk stratification in preoperative settings[10]. Xue and colleagues created five machine-learning models to predict lung complications after gastrointestinal surgery. Their logistic regression model had an AUC of 0.808 with an accuracy of 0.824 and a precision of 0.621. The decision tree model had an AUC of 0.702 with an accuracy of 0.795 and a precision of 0.486 and the gradient boosting model obtained an AUC of 0.814 with an accuracy of 0.806 and precision of 0.750. The other two models were comparable[11]. Moreover, retrospective data from previous surgeries, such as the clinical decisions made perioperatively by the anesthesiologists and the postoperative functional outcomes, could teach AI to give recommendations[10].

#### *Airway difficulties*

AI is used in pediatric patients to predict difficult airways by analyzing 2D and 3D facial features, with a real-time application to detect tracheal anatomy, based on data from bronchoscopies[12].

An estimated 75-93% of the problematic intubations are unanticipated[13]. Jong Ho Kim et al. conducted a retrospective cohort study on Asian subjects. They created an AI model to predict difficult intubation using only three variables: Mallampati grade, age and sternomental distance. Using a random forest prediction model, their model achieved an area under the receiver operating curve of 0.71 (95% CI 0.72 - 0.86)[14].

Hayasaka et al. created a neural network model to link facial characteristics with airway difficulties. They took pictures of 202 patients in 16 different positions. Their model achieved a high predictive value with an area under the curve of 0.864 (95% CI 0.731 - 0.969) with an accuracy of 80.5%, sensitivity of 81.8%, specificity of 83.3%, based on the picture taken in a supine neutral position with mouth closed[15].

#### *Operating theatre logistics*

Hashimoto et al. reviewed three studies that used AI for operating room logistics, but none achieved more than 60% accuracy in predicting surgical duration[16,17]. AI could improve OR management, workflow, and scheduling to increase efficiency[9].

### *Intraoperative use of AI*

Intraoperative awareness is a dreadful complication, as is prolonged hypotension, massive bleeding or residual curarization. Thanks to artificial intelligence, aids have been developed to predict and detect these events. However, with yet another tool warning us, the alarm overload could cause inattention or alarm fatigue.

Intraoperative imaging is becoming state-of-the-art clinical perioperative practice. However, it is user dependent and has a steep learning curve. Two AI applications increasing the accuracy of perioperative ultrasound are discussed.

#### *Depth of anesthesia*

AI facilitates the ultimate development of closed-loop systems by considering individual patient variability when controlling the delivery of anesthesia[18].

Joosten et al. showed in a randomized controlled trial that automated anesthetic management using the combination of three controllers (end-tidal CO2 concentration between 32 and 38 mmHg, fluid balance through continuous infusion and boluses of 100 mL, based on analyses of stroke volume, heart rate, mean arterial pressure, stroke volume variation, and BIS > 40) outperformed manual control. People in the closed loop group showed a better cognition score one week after surgery. This effect persisted three months after surgery[19].

Tacke et al. wanted to detect awareness using AI technology. However, they only collected data from EEG and auditory evoked potentials before and after, but not during surgery, rendering a final evaluation of the ability of AI to discriminate between perioperative consciousness and unconsciousness complex[20].

Control systems for infusions of neuromuscular blocking drugs or (weaning from) mechanical ventilation have been described[9]. Neural networks have been used to predict recovery from neuromuscular blockade and hypotension after induction or spinal anesthesia[9].

Data (abdominal girth, vertebral column length) of 684 parturients undergoing cesarean section were used for a machine-learning algorithm to identify the optimal dose of intrathecal 0.5% hyperbaric bupivacaine to achieve an appropriate block level from T4-T6 without risking hypotension[21]. Wei et al., therefore, developed a decision model with a determination coefficient ($R^2$) of 0.8070 and a mean squared error (MSE) of 0.0087, meaning the model is nearly perfect[21].

*Predicting adverse events*

Yet undiagnosed heart failure with reduced ejection fraction was detected in 0.41% of 67697 patients undergoing non-cardiac surgery via machine learning approaches. The AUC for their logistic regression model was 0.869 (95% CI 0.829 - 0.911), 0.872 (95% CI 0.836 - 0.909) for the random forest model and 0.873 (95% CI 0.833 - 0.913) for the extreme gradient boosting model. However, the low prevalence of the disease in this population resulted in a low positive predictive value, mandating confirmatory testing with high specificity[22].

Bai et al. developed an ML method to analyze cerebral and myocardial infarction risk factors after carotid endarterectomy. Studying 443 patients, the incidence of cerebral infarction was 1.4%, and that of myocardial infarction was 2.3%. They identified eight predictive factors such as blood pressure, body mass index, and age. However, in cross-validation, their prediction model was highly fluctuating[23].

Blood loss during surgery can be challenging to estimate. Through image acquisition and connected convolutional networks, estimated blood loss and estimated hemoglobin loss were estimated more accurately. They obtained an R2 of 0.966 (95% CI 0.962 – 0.971)[24].

Artificial intelligence (AI) is increasingly used to predict intraoperative hypotension (IOH)[25]. Kendale et al. compared several predictive models and their AUC, namely: logistic regression, support vector machines, naïve Bayes, k-nearest neighbor, linear discriminant analysis, random forest, neural nets and gradient boosting machine. The AUC of the gradient boosting machine was the highest, namely 0.76 (95% CI, 0.75 - 0.77). The test set AUC for the gradient boosting machine was 0.74 (95% CI, 0.72 - 0.77)[25].

Lee et al. developed a random forest model that predicts hypotension with 74.89% accuracy one minute in advance[26]. Frassanito et al. used continuous arterial pressure waveforms to predict IOH with an AUC of 0.93 (95% CI 0.89 - 0.97), 0.9 (95% CI 0.83 - 0.97), and 0.95 (95% CI 0.89 - 0.99) respectively 5, 10, and 15 minutes before the event[27].

Li et al. obtained an AUC of 0.843 (95% CI 0.808 - 0.877) in predicting hypotension in cardiac surgery using a random forest model[28]. AI models such as HemoSphere (Edwards Lifesciences, Irvine, Washington, United States) are commercially available for hypotension prediction[29]. In a study by Hollman, using hypotension prediction models, this group experienced less perioperative hypotension than compared to standard care[30].

*Monitoring and alarm settings*

Conway tried to create a “smart alarm” that can alert prior to a predicted prolonged apneic event (i.e. > 30 seconds) in procedural sedation, using capnography. However, this model was not superior to the conservative strategy[31].

*Perioperative imaging*

Echocardiographical images are frequently associated with interobserver variability[32]. The “eSie Valve Software” (Siemens, Groot-Bijgaarden, België) was tested in the retrospective analysis of perioperative 3D TEE data from four patients undergoing CABG and found to reliably perform automated analyses of the mitral valve with good reproducibility[32].

AutoLV (TomTec-Arena, Unterschleissheim, Germany) can offer automated left ventricle ejection fraction measurements in only 8 seconds. It takes 82% less time and correlates better with cardiac output measured by thermodilution than by manual echocardiographic measurements[33].

***Postoperative use of AI***

The opioid crisis raised awareness about the use of opioids. In chronically opioid-exposed patients, pain management after surgery can be challenging. Two studies investigated this subject.

Postoperatively, it is crucial not only to predict pain but also postoperative organ dysfunction, delirium and in-hospital mortality. AI predictive models have been described for each of these subjects.

*Pain management*

AI has been used to predict opioid dosing and which patients would respond to opioid therapy for acute pain; based on preoperative electroencephalography assessment[9]. In children, Cheng et al. proposed a study protocol of a systematic review and meta-analysis to compare different facial expression-based ML algorithms to evaluate pain[34].

*Prediction of postoperative organ dysfunction or mortality*

Zhang et al. retrospectively analyzed 780 adult liver transplant cases, creating a gradient-boosting model with an AUC for the internal validation set of 0.76 (95% CI 0.70 - 0.82) and an AUC of 0.75 (95% CI 0.67 - 0.81) on the external validation set to predict acute kidney injury (AKI)[31]. Lee et al. found that their gradient-boosting technique showed better accuracy than logistic regression in predicting new-onset AKI in a retrospective study of 2010 patients, respectively an AUC of 0.78 (95% 0.75 - 0.80) versus an AUC of 0.69 (95% CI 0.66 - 0.72)[36].

Zhao et al. used ML methods to detect 16 risk factors for postoperative delirium in older patients undergoing hip fracture surgery, achieving an AUC of 0.779 (95% CI 0.703 – 0.856). Preparation time, frailty index, use of vasopressors during the surgery, dementia or history of stroke, duration of surgery, and anesthesia were the six most important predictors of delirium[37].

Lee et al. also investigated ML models to predict postoperative in-hospital mortality using data from 59 985 surgical records. Their GAM Neural network model achieved an AUC of 0.921 (95% CI 0.895 – 0.95)[38].

### *AI in the intensive care unit*

Predicting outcomes and mortality of patients admitted at an ICU can be highly challenging. By analyzing large datasets, different investigators' groups created an AI model improving the current state of the art. AI tools have been developed as decision aid systems or as a diagnostic aid to manage life-threatening situations with often poor outcomes. AI can also be incorporated into less critical circumstances that improve patient care and well-being, like pain management, alarm settings or dental hygiene.

#### *Predicting outcome and mortality*

At the ICU, AI has been proven helpful in severity scoring, mortality prediction and prediction of sepsis in an early phase[8]. AI outperformed systemic inflammatory response syndrome criteria (SIRS) and the sequential organ failure assessment score (SOFA score) in detecting early sepsis, decreasing hospital length of stay and in-hospital mortality[9]. This is confirmed by Churpek et al. during the COVID-19 pandemic. Including 5075 patients in 68 different United States ICUs, their extreme gradient boosting model had an AUC of 0.81 (95% CI 0.78 - 0.85) compared to 0.69 (95% CI 0.65 - 0.73) from the SOFA score or AUC of 0.60 (95% CI 0.55 - 0.64) from the National Early Warning Score, in predicting in-hospital death within 28 days of the ICU admission[39].

AI has also been used to predict neurological outcomes after out-of-hospital cardiac arrest (OHCA). Andersson et al. performed a post hoc multicenter analysis of 932 patients. Their model based on integrated clinical variables and accessible biomarkers such as NSE (neuron-specific enolase) achieved an AUC of 0.94 (95% CI 0.894 - 0.988). They aimed for zero false-positive predictions, which could result in the withdrawal of life-sustaining therapies in patients who would have had a good outcome[40].

Another European investigation was conducted using neural networks to predict outcomes in OHCA patients. With only three variables (age, time to return of spontaneous circulation (ROSC) and first monitored rhythm), their ML model showed an AUC above 0.852 (95% CI 0.835 – 0.869) in predicting 180 days functional outcome, including survival. When 54 variables were considered, the AUC rose even more but is not practical in its use[41].

Ghassemi and his team analyzed 12 397 hours of EEG from 438 subjects to predict recovery in hypoxic-ischemic encephalopathy. They obtained an AUC of 0.83 (95% CI 0.75 – 0.91) in predicting 6-month functional outcomes[42].

#### *Diagnostic aid and decision making*

ARDS is difficult to discriminate from acute cardiogenic pulmonary oedema (CPE), especially considering the recent pandemic[43]. Therefore, Brusasco and her team developed a computer-aided diagnosis to determine those two life-threatening conditions on pleural ultrasound images, including patients prospectively. Their model was based on different texture features seen on ultrasound; for example, the AUC for contrast-based textural differences was 0.891 (95% CI 0.726 – 1.000) in differentiating both diseases[44].

Not only in pulmonary diseases but also in sepsis, AI has proven its use. Komorowski and colleagues developed the AI Clinician, a computational model using reinforcement learning, which can dynamically suggest optimal treatments (administration of fluid or vasopressors) for adult patients with sepsis in the intensive care unit (ICU). In an independent cohort, the patients who received the treatments suggested by the AI Clinician had the lowest mortality rate[45].

In Hong Kong, researchers developed a deep learning model using data from 10941 critically ill patients from 209 ICUs to predict the need for vasopressor therapy within the first two hours of ICU admission. This neural network-based Bi-LSTM model achieved an area under the curve of 0.96 (95%CI 0.96 - 0.96). Heart rate, respiratory rate, and mean arterial pressure contributed most to the model, amongst other serial physiological variables of systolic blood pressure, diastolic blood pressure, pulse oximetry and temperature[46].

#### *Monitoring pain*

A team from Singapore analyzed 746 video clips of 63 critically ill patients using AI. Although they experienced difficulties since medical devices and oedema masked some facial areas, their deep-learning-based pain classifier did detect pain with an accuracy of 80%-90%[47].

*Ventilation*

In pandemic situations, real-time monitoring can be challenging. Radhakrishnan et al. developed a neural network model to predict the level of inspired oxygen delivered by the mechanical ventilator along with ventilation mode and positive end-expiratory pressure (PEEP) changes to reduce the efforts of healthcare professionals[48].

*Alarm settings*

The high number of false positive alarms at the ICU increases the workload for the personnel. AI could help to reduce alarm fatigue[49]. Fernandes et al. state that 80-99% of alarms in hospital units are false or clinically insignificant, leading to an alert overload. Providing a reasoning system based on AI could reduce the notifications by up to 99.3%[50].

*Day-to-day care*

Scquizzato introduced smart toothbrushes to help prevent ventilator-associated pneumonia in critically ill patients, using built-in 3D[51].

## *AI as a tool for teaching*

Simulation classes prepare young colleagues for adverse events in a safe learning environment. Many articles investigated this subject with positive results.

*Neuraxial and locoregional anesthesia*

Chan et al. developed an ultrasound-guided automated spinal landmark identification program to identify the insertion point of the spinal needle. In 48 obese patients, the first-attempt success rate was 79.1%. This software can also be helpful in patients with abnormal spine, scoliosis, and previous spinal surgery. In this population, they reported a first-attempt success of 32% with palpation and an improvement to 65% using ultrasound imaging[52].

In epidural procedures in obstetric anesthesia, machine learning has proven faster and more dose-effective in achieving a sensory level. The AI framework can also predict the incidence of hypotension with 85% accuracy[53].

Combining ultrasound and AI with medical image fusion while performing neuraxial blocks can improve the success rate in patients with obesity, elderly patients and trauma patients with tissue oedema or fat infiltration and train residents. One main problem is that such a system cannot effectively avoid intraneural injection yet[54].

Gungor et al. assessed the accuracy of an AI-based real-time anatomy identification software specifically designed for peripheral nerve block. Forty healthy participants underwent four ultrasound procedures (supraclavicular, infraclavicular, inter scalene and transversus abdominis plane block (TAP)) without block performance. They conducted a kappa test to examine the agreement level between the two expert validators' evaluation scores. The Cohen's kappa value for the TAP block, inter scalene block, supraclavicular and infraclavicular block were respectively 0.95, 0.98, 0.96 and 0.97, indicating that AI technology can successfully interpret anatomical structures in real-time sonography while assisting young anesthesiologists during practice[55].

AnatomyGuide (Intelligent Ultrasound Limited, Cardiff, UK) is a system based on AI technologies. One hundred twenty thousand images were used in the training set for each block. The initial skill acquisition and the period required for direct supervision could be shortened by highlighting the relevant structures. However, this technology is expensive and does not indicate nuances such as probe pressure, angulation, rotation, tilting or needle-probe coordination[56].

Another AI model improved the accuracy of the image and reduced the time from needle puncture to completion of injection (7.5 minutes compared to 10.2 minutes in the control group) in 100 patients with a scapular fracture to perform nerve block[57].

*Simulation education*

Using AI, fictional case scenarios can be generated. For example, artificial breast ultrasound images have been created based on a pre-existing database. Generating these artificial images can be a time-gaining step in preparing scenarios[58].

*Critical remarks on AI implementation in healthcare*

Artificial intelligence became very popular in a short period of time. It is often presented as the solution to many problems, the holy grail, without realizing this technology can also cause some big issues. Geoffrey Hinton, the man often referred to as the "godfather of AI", recently quit his job at Google and warned us in multiple statements for the possible dangers of this growing technology.

*Data-dependency*

AI can aid human providers in reducing the work burden and compiling and analyzing data but cannot replace human beings in day-to-day care[59]. The big obstacle to the widespread implementation of AI in healthcare is the mandatory access to large amounts of good quality data, upon which AI algorithms are built. Data protection issues play a major role in this context[6]. Better, more comprehensive and easily accessible datasets are prerequisites for AI applications to become clinically meaningful[9]. Data from different types of populations, races, or extremes of age should be incorporated into these machine-learning programs to avoid bias. There is no learning beside the given data possible.

*Monitoring artefacts*

Many of the AI strategies are based on data derived from (non-)invasive monitoring in the OR, which is vulnerable to various artefacts. AI must correct its underlying algorithms for this type of 'noise' or alert the physician that something is wrong with the input signal[60]. For example, in acceleromyographic neuromuscular monitoring, AI successfully reduced outliers and increased reliability[61].

*The role of the anesthesiologist*

With all these new technologies rapidly developing, will anesthesiologists become unnecessary one day? While AI could protect the human provider from cognitive overload, some authors have suggested that the anesthesiologists' "hands" will always remain necessary, making full automation impossible[62]. However, if anesthesia is reduced to its mere technicity, an essential asset of our work could be lost.

If AI is to be embraced as a decision-making tool or clinical support system, healthcare providers can focus on "higher-order" clinical decision-making and patient care[10]. However, caregivers should know about the variables on which AI algorithms are based and whether and how they can overrule machine learning if necessary. Therefore, anesthesiologists should partner with engineers and computer specialists to avoid the well-known problem of the "black box" and gain full transparency into how clinical decisions are made. Furthermore, this multidisciplinary approach should involve data scientists, ethicists, project managers, analysts and end users[63].

Last, the main goal for the implementation of AI is to reduce workload and simplify different data streams. However, with just another machine in the OR, it must be prevented AI itself causes data overload.

*Legal and ethical framework*

Before new AI strategies are implied in daily practice, accountability should be defined: the software engineers or the physician using this new type of technology.

Specialists in medical ethics will have to discuss whether it is ethical to partially or even entirely rely on autonomous thinking machines and their predictions[64].
We must ensure that human interaction will not be lost and that our patients are given the empathy and human reassurance they need[65].

The question arises how the certification of this software will be granted and how this certification can be renewed when the software is updated with new data gained from clinical exposure.

*Costs*

Not a single article mentioned the purchase price or maintenance costs of the software program tested and the devices needed to implement it in clinical practice.

We found that AI technology will trigger an estimated $147 billion market during the next 20 years and hereby transform the medical field[66]. Although AI could be a solution to (partially) face the problems caused by the increasing workforce shortage and economization, this technology on its own will be associated with considerable costs.

*Safety*

Several concerns about the "black box" within the AI systems have been brought forward[67]. The way machines learn is not transparent at all. However, it is vital for healthcare providers to know why AI took a certain decision and how interference is possible[68].

AI reasoning, without clinicians knowing the variables on which certain predictions are based, is associated with important risks of safety.
Likewise, cybersecurity is also of utmost importance to prevent hacking of AI-based medical applications with potentially deadly consequences.

## Discussion

In this review, we discussed the omnipresence of AI in medicine. We gave an overview and definitions of frequently used AI terminology. Secondly, we gave an overview of previous research within the field of anesthesiology and categorized our findings with regard to applications in pre-, intra- and postoperative care, intensive care and teaching. These categories correspond to the relevant steps in the anesthetic workflow.

Our findings are similar to other published work. Two papers were issued with a scope corresponding with ours[8,9]. However, the strength of the present paper is not only to frame AI-related definitions, but also to summarize and categorize current publications and offer critical remarks and concerns. Moreover, we included very recent articles, serving as an update on previously published reviews.

Artificial intelligence is rapidly evolving as an engineering discipline, but it is crucial for medical professionals to ensure these technical possibilities can be implemented in a medically responsible way. Continuous research is important. The New England Journal of Medicine therefore announced the launch of their new journal NEJM AI in 2024, aiming to provide high-quality evidence for medical AI along with informed discussions[69].

### *Limitations*

This paper has some limitations. Firstly, as it is not a systematic review, articles may have been missed. Secondly, the included papers are only briefly discussed, since this study aimed to give an overview of existing literature and applications, rather than discussing each topic in depth. Thirdly, we acknowledge that by the time this article is published, new insights may have been acquainted so that the present paper can be outdated quickly because of the rapidly evolving AI world.

## Conclusion

John McCarthy, the father of "artificial intelligence", once stated: "As soon as it works, no one calls it AI anymore". Probably, AI will be as normal as traditional monitoring, such as a pulse oximeter which we routinely use without thinking about the Lambert-Beer law every single time.

While many opportunities for AI are yet to come, expectations should not become unrealistic. AI is certainly not the answer to all problems in healthcare.

One of the primary goals in implementing a new technology in medicine is to reduce human error and to improve patient safety. While AI offers numerous possibilities, AI is critically dependent on large, accurate datasets. Firstly, the anesthesiologist community can help to build these datasets, and hereby create a legal, ethical and easily accessible framework for engineers for study purposes. Partnering with software engineers is necessary. Most importantly, AI applications must be prospectively tested in multicenter trials since most of the current research is retrospective and derived from single center evidence.

Even when AI can help to decrease healthcare costs and improve efficacy, it should be remembered that AI has a price tag on its own as well. Studies on cost-effectiveness are warranted.

Careful implementation and use of AI, along with real-time human interpretation, will revolutionize perioperative care and is the way forward in future perioperative management of major surgery[70]. As we are going towards "smart healthcare", anesthesiologists will have to reinvent themselves to make sure the art of anesthesia will not be lost and still be able to perform anesthesia without relying on technical tools, not always available in critical situation [71].

*Acknowledgments:* I would like to thank Professor doctor Steffen Rex for his help and advice writing this paper.

*Conflict of interests:* None.

## References

1. Kung TH, Cheatham M, Medenilla A, Sillos C, Leon L De, Elepa C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digit Heal. 2023;1–12.
2. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. JAMA. 2023;1–8.
3. Seger C, Cannesson M. Recent advances in the technology of anesthesia. F1000 Research [online serial] 2020; https://doi.org/10.12688/f1000research.24059.1 [accessed 18 April 2022].
4. Zippel C, Bohnet-joschko S. Rise of Clinical Studies in the Field of Machine Learning: A Review of Data Registered in ClinicalTrials.gov. Int J Environ Res Public Health 2020: 18(10).
5. Alexander JC, Joshi GP. Anesthesiology, automation, and artificial intelligence. Proc Bayl Univ Med Cent 2018; 31(1): 117-119.
6. Mc Grath H, Flanagan C, Zeng L, Lei. Anaesthesia Monitoring using Artificial Intelligence Techniques. Int J Anesthetic Anesthesiol [online serial] 2019; 6:098. doi.org/10.23937/2377-4630/1410098 [accessed 27 March 2023]
7. Yongjun Xu, Xin liu, Xin Cao e.a. Artificial intelligence : A powerful paradigm for scientific research. Innov. [online serial] 2021; https://doi.org/10.1016/j.xinn.2021.100179 [accessed 2 September 2022].
8. Naidoo S. Understanding Artificial Intelligence and Machine Learning in Anaesthesia. 2021;(01):1–17. [accessed 20 September 2022].
9. Hashimoto, D. A., et al. Artificial Intelligence in Anesthesiology: Current Techniques, Clinical Applications, and Limitations. Anesthesiology. 2020; 132(2): 379-394.
10. Côté, Clémence D. and Paul J. Kim. Artificial intelligence in anesthesiology: Moving into the future. UTMJ. 2019; 158-162.
11. Xue Q, Wen D, Ji M, Tong J, Yang J. Developing Machine Learning Algorithms to Predict Pulmonary Complications After Emergency Gastrointestinal Surgery. Front Med (Lausanne) 2021; 8:1–6.
12. Matava, C., et al. Artificial intelligence, machine learning and the pediatric airway. Paediatr Anaesth 2020; 30(3): 264-268.
13. Tavolara, T. E., et al. Identification of difficult to intubate patients from frontal face images using an ensemble of deep learning models. Comput Biol Med 2021; 136: 104737.
14. Kim, J. H., 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.
15. Hayasaka, T., et al. 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.
16. Devi, S. P., et al. Prediction of surgery times and scheduling of operation theaters in ophthalmology department. J Med Syst 2012; 36(2): 415-430.
17. Houliston BR, Parry DT, Merry AF. TADAA: Towards Automated Detection of Anaesthetic Activity. Methods Inf Med. 2011; 50(5):464-71. doi: 10.3414/ME11-02-0001.
18. LE Guen M, Liu N, Chazot T, Fischler M. Closed-loop anesthesia. Minerva Anestesiol. 2016; May; 82(5):573-81. PMID: 26554614.
19. Joosten A, Rinehart J, e.a. Anesthetic Management Using Multiple Closed-loop Systems and Delayed Neurocognitive Recovery: A Randomized Controlled Trial. Anesthesiology 2020; Feb;132(2):253-266. doi: 10.1097/ALN.0000000000003014.
20. Tacke, M., et al. Machine learning for a combined electroencephalographic anesthesia index to detect awareness under anesthesia. PLoS One 2020 15(8): e0238249.
21. Wei, C. N., et al. A prediction model using machine-learning algorithm for assessing intrathecal hyperbaric bupivacaine dose during cesarean section. BMC Anesthesiol 2021; 21(1): 116.
22. Mathis MR, Engoren MC, e.a. Early Detection of Heart Failure With Reduced Ejection Fraction Using Perioperative Data Among Noncardiac Surgical Patients: A Machine-Learning Approach. Anesth Analg. 2020; May;130(5):1188-1200. doi: 10.1213/ANE.0000000000004630.
23. Bai, P., et al. Risk Factors of Cerebral Infarction and Myocardial Infarction after Carotid Endarterectomy Analyzed by Machine Learning. Comput Math Methods Med 2020; e6217392.
24. Li Y, Zhang L, Zhi H, Zhong K, He W, Chen Y, et al. A better method for the dynamic , precise estimating of blood / haemoglobin loss based on deep learning of artificial intelligence. Ann Transl Med 2020; 8(19): http://dx.doi.org/10.21037/atm-20-1806 [accessed 2 November 2022]
25. Kendale, S., et al. Supervised Machine-learning Predictive Analytics for Prediction of Postinduction Hypotension. Anesthesiology 2018; 129(4): 675-688.
26. Lee, J., et al. Comparative Analysis on Machine Learning and Deep Learning to Predict Post-Induction Hypotension. Sensors (Basel) 2020; 20(16).
27. Frassanito L, Paolo P, Francesco G, Alessandra V, Alessia P, Antonio B, et al. Hypotension Prediction Index with non invasive continuous arterial pressure waveforms ( ClearSight ): clinical performance in Gynaecologic Oncologic Surgery. J Clin Monit Comput [Internet]. 2021; https://doi.org/10.1007/s10877-021-00763-4.
28. Li X-F, Huang Y-Z, Tang J-Y. Development of a random forest model for hypotension prediction after anesthesia induction for cardiac surgery. World J Clin cases. 2021; 9(29): 8729–40.
29. Javier G, Daniel S, Manuel G, Monge I, Carlos G, Jose C, et al. Real world outcomes of the hypotension prediction index in the management of intraoperative hypotension during non cardiac surgery: a retrospective clinical study. J Clin Monit Comput [Internet]. 2022; https://doi.org/10.1007/s10877-022-00881-7.
30. Hollmann MW, Vlaar AP, Veelo DP. 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.
31. Conway, A., et al. Predicting Prolonged Apnea During Nurse-Administered Procedural Sedation: Machine Learning Study. JMIR Perioper Med 2021; 4(2): e29200.
32. Jeganathan, J., et al. Artificial intelligence in mitral valve analysis. Ann Card Anaesth 2017; 20(2): 129-134.
33. Chen X, Owen CA, Huang EC, Maggard BD, Latif RK, Clifford SP, et al. Artificial Intelligence in Echocardiography for Anesthesiologists. J Cardiothorac Vasc Anesth [Internet]. 2021; 35(1):251–61. https://doi.org/10.1053/j.jvca.2020.08.048.
34. Cheng, D., et al. Current state of science in machine learning methods for automatic infant pain evaluation using facial expression information: study protocol of a systematic review and meta-analysis. BMJ Open 2019; 9(12): e030482.
35. Zhang Y, Yang D, Liu Z, Chen C, Ge M, Li X, et al. An explainable supervised machine learning predictor of acute kidney injury after adult deceased donor liver transplantation. J Transl Med [Internet]. 2021; 1–15. https://doi.org/10.1186/s12967-021-02990-4.
36. Lee, H. C., et al. Derivation and Validation of Machine Learning Approaches to Predict Acute Kidney Injury after Cardiac Surgery. J Clin Med 2018; 7(10).
37. Zhao, H., et al. Machine Learning Algorithm Using Electronic Chart-Derived Data to Predict Delirium After Elderly Hip Fracture Surgeries: A Retrospective Case-Control Study. Front Surg 2021; 8: 634629.
38. Lee CK, Samad M, Hofer I, Cannesson M. Development and validation of an interpretable neural network for prediction of postoperative in-hospital mortality. npj Digit Med [Internet]. http://dx.doi.org/10.1038/s41746-020-00377-1.
39. Churpek, M. M., et al. Machine Learning Prediction of Death in Critically Ill Patients With Coronavirus Disease 2019. Crit Care Explor 2021; 3(8): e0515.
40. Andersson P, Johnsson J, Björnsson O, Cronberg T, Hassager C, Zetterberg H, et al. Predicting neurological outcome after out of hospital cardiac arrest with cumulative information ; development and internal validation of an artificial neural network algorithm. Crit Care [Internet]. 2021;1–12. https://doi.org/10.1186/s13054-021-03505-9.
41. Johnsson, J., et al. Artificial neural networks improve early outcome prediction and risk classification in out-of-hospital cardiac arrest patients admitted to intensive care. Crit Care 2020; 24(1): 474.
42. Ghassemi M, Herman ST, Sivaraju A, Gaspard N, Junior M, Cash SS, et al. Quantitative EEG trends predict recovery in hypoxic-ischemic encephalopathy. Crit Care Med 2019; 47(10): 1416-1423
43. Mushtaq, J., et al. Initial chest radiographs and artificial intelligence (AI) predict clinical outcomes in COVID-19 patients: analysis of 697 Italian patients. Eur Radiol 2021; 31(3): 1770-1779.
44. Brusasco C, Santori G, Tavazzi G, Via G, Robba C, Gargani L, et al. Second order grey scale texture analysis of pleural ultrasound images to differentiate acute respiratory distress syndrome and cardiogenic pulmonary edema. J Clin Monit Comput [Internet]. 2020;(0123456789). https://doi.org/10.1007/s10877-020-00629-1.
45. Komorowski, M., et al. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 2018; 24(11): 1716-1720.
46. Kwak GH. Predicting Need for Vasopressors in the Intensive Care Unit Using an Attention Based Deep Learning Model. SHOCK. 2021;56(1):73–79.
47. Wu, C. L., et al. Deep Learning-Based Pain Classifier Based on the Facial Expression in Critically Ill Patients. Front Med (Lausanne) 2022; 9: 851690.
48. Radhakrishnan, S., et al. Multilayer perceptron neural network model development for mechanical ventilator parameters prediction by real time system learning. Biomed Signal Process Control 2022 71: 103170.
49. Poncette, A. S., et al. Improvements in Patient Monitoring in the Intensive Care Unit: Survey Study. J Med Internet Res 2020; 22(6): e19091. [accessed 15 December 2022]
50. Fernandes CO, Miles S, Jos C, Cowan D. Artificial Intelligence Technologies for Coping with Alarm Fatigue in Hospital Environments Because of Sensory Overload : Algorithm Development and Validation. J Med Internet Res. 2019;21(11).

51. Scquizzato, T. and A. Gazzato. Adopting a smart toothbrush with artificial intelligence may improve oral care in patients admitted to the intensive care unit. Crit Care 2019: 23(1): 223.
52. Chan I, Anesthesiol BMC, Ju J, Chan I, Ma J, Leng Y, et al. Machine learning approach to needle insertion site identification for spinal anesthesia in obese patients. BMC Anesthesiol [Internet]. 2021;1–8. https://doi.org/10.1186/s12871-021-01466-8.
53. Lee, Eva K.; Tian, Haozheng; Lee, Jinha. Investigating a Needle-Based Epidural Procedure in Obstetric Anesthesia. Public and Allied Health Faculty Publications 2018. https://scholarworks.bgsu.edu/publ_allied_health_pub/35.
54. Wu Z, Wang Y. Development of Guidance Techniques for Regional Anesthesia: Past, Present and Future. J Pain Res. 2021 Jun 9;14:1631-1641. doi: 10.2147/JPR.S316743.
55. Gungor I, Gunaydin B, Oktar SO, Buyukgebiz BM, Bagcaz S, Gozde M. A real time anatomy ıdentification via tool based on artificial ıntelligence for ultrasound guided peripheral nerve block procedures: an accuracy study. J Anesth [Internet]. 2021;35(4):591–4. https://doi.org/10.1007/s00540-021-02947-3.
56. Bowness J. Artificial intelligence for image interpretation in ultrasound-guided regional anaesthesia. Anaesthesia. 2021;76:602–7.
57. Liu, Y. and L. Cheng. Ultrasound Images Guided under Deep Learning in the Anesthesia Effect of the Regional Nerve Block on Scapular Fracture Surgery. J Healthc Eng 2021: 6231116.
58. Arora A. Artificial intelligence : a new frontier for anaesthesiology training Would you trust your loved ones to this trainee? Certification decisions in postgraduate anaesthesia training. Br J Anaesth [Internet]. 2020;125(5):e407–8. https://doi.org/10.1016/j.bja.2020.06.049.
59. Garg R, Patel A, Hoda W. Emerging role of artificial intelligence in medical sciences-Are we ready! J Anaesthesiol Clin Pharmacol. 2021; 37(1):35-36. doi: 10.4103/joacp.JOACP_634_20.
60. Pasma W, Wesselink EM, Buuren S Van, Graaff JC De, Klei WA Van. Artifacts annotations in anesthesia blood pressure data by man and machine. J Clin Monit Comput [Internet]. 2021;35(2):259–67. https://doi.org/10.1007/s10877-020-00574-z.
61. Carvalho H, Berghmans J, Forget P. Exploratory Outlier Detection for Acceleromyographic Neuromuscular Monitoring: Machine Learning Approach. J Med internet Res. 2021;23(6):1–12.
62. Char DS, Burgart A, Food R, Kettering S. Machine learning implementation in clinical anesthesia: opportunities and challenges. Anesth Analg. 2020;130(6):1709–1712.
63. Lonsdale, H., et al. Artificial Intelligence in Anesthesiology: Hype, Hope, and Hurdles. Anesth Analg 2020; 130(5): 1111-1113.
64. Abramoff, M. D., et al. Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process. Am J Ophthalmol 2020; 214: 134-142.
65. Hofer, I. S., et al. Realistically Integrating Machine Learning Into Clinical Practice: A Road Map of Opportunities, Challenges, and a Potential Future. Anesth Analg 2020: 130(5): 1115-1118.
66. Kong X, Ai B, Kong Y, e.a. Artificial intelligence: a key to relieve China's insufficient and unequally-distributed medical resources. Am J Transl Res. 2019 May 15;11(5):2632-2640. PMID: 31217843; PMCID: PMC6556644. [accessed 14 December 2022]
67. Gambus PL, Jaramillo S. Machine learning in anaesthesia: reactive , proactive … predictive! Br J Anaesth [Internet]. 2019;123(4):401–3. https://doi.org/10.1016/j.bja.2019.07.009 [accessed 14 December 2022]
68. Mathis, M. R., et al. Artificial Intelligence for Anesthesia: What the Practicing Clinician Needs to Know: More than Black Magic for the Art of the Dark. Anesthesiology 2018; 129(4): 619-622.
69. Leong TY, Ph D, Manrai AK, Ph D, Rubin EJ, Ph D. Artificial Intelligence in Medicine. N Engl J Med. 2023;388(13).
70. Solanki, S. L., et al. Artificial intelligence in perioperative management of major gastrointestinal surgeries. World J Gastroenterol 2021; 27(21): 2758-2770.
71. Hung O. Artificial intelligence versus the art of anesthesia: a long and winding road ahead [Internet]. Can J Anaesth 2023. https://doi.org/10.1007/s12630-023-02473-8.

doi.org/10.56126/75.3.21