



REVIEW

Open Access

Machine learning in perioperative medicine: a systematic review



Valentina Bellini¹, Marina Valente², Giorgia Bertorelli¹, Barbara Pifferi¹, Michelangelo Craca¹, Monica Mordonini³, Gianfranco Lombardo³, Eleonora Bottani³, Paolo Del Rio² and Elena Bignami^{1*}

Abstract

Background: Risk stratification plays a central role in anesthetic evaluation. The use of Big Data and machine learning (ML) offers considerable advantages for collection and evaluation of large amounts of complex health-care data. We conducted a systematic review to understand the role of ML in the development of predictive post-surgical outcome models and risk stratification.

Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines, we selected the period of the research for studies from 1 January 2015 up to 30 March 2021. A systematic search in Scopus, CINAHL, the Cochrane Library, PubMed, and MeSH databases was performed; the strings of research included different combinations of keywords: “risk prediction,” “surgery,” “machine learning,” “intensive care unit (ICU),” and “anesthesia” “perioperative.” We identified 36 eligible studies. This study evaluates the quality of reporting of prediction models using the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) checklist.

Results: The most considered outcomes were mortality risk, systemic complications (pulmonary, cardiovascular, acute kidney injury (AKI), etc.), ICU admission, anesthesiologic risk and prolonged length of hospital stay. Not all the study completely followed the TRIPOD checklist, but the quality was overall acceptable with 75% of studies (Rev #2, comm #minor issue) showing an adherence rate to TRIPOD more than 60%. The most frequently used algorithms were gradient boosting ($n = 13$), random forest ($n = 10$), logistic regression (LR; $n = 7$), artificial neural networks (ANNs; $n = 6$), and support vector machines (SVM; $n = 6$). Models with best performance were random forest and gradient boosting, with AUC > 0.90.

Conclusions: The application of ML in medicine appears to have a great potential. From our analysis, depending on the input features considered and on the specific prediction task, ML algorithms seem effective in outcomes prediction more accurately than validated prognostic scores and traditional statistics. Thus, our review encourages the healthcare domain and artificial intelligence (AI) developers to adopt an interdisciplinary and systemic approach to evaluate the overall impact of AI on perioperative risk assessment and on further health care settings as well.

Keywords: Risk prediction, Surgery, Machine learning, ICU, Anesthesia, Perioperative

* Correspondence: elenagiovanna.bignami@unipr.it

¹Anesthesiology, Critical Care and Pain Medicine Division, Department of Medicine and Surgery, University of Parma, Viale Gramsci 14, 43126 Parma, Italy

Full list of author information is available at the end of the article



© The Author(s). 2022 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

Risk stratification is a central part of the anesthetic evaluation. In fact, through the identification of high-risk patients, it is possible to conduct a specific risk/benefit analysis, to reduce the risk of unexpected complications, to achieve a targeted perioperative optimization, to carefully plan the anesthesiologic management, and to provide an accurate and precise informed consent [1–3].

Over time, several scores have been published, from the most generic, like the American Society of Anesthesiologists Physical Status (ASA-PS) [4], to the most specific ones, as the European system for cardiac operative risk evaluation (EuroSCORE) [5] or the General Surgery Acute Kidney Injury Risk Index Classification System [6]. Unfortunately, these scores have some limits, mainly due to the lack of tailored predictions.

In the last decade, the interest about artificial intelligence (AI), including machine learning (ML) methods, have seen an exponential increase [2]. Considered an extension of traditional statistics, AI differs from standard approaches for its ability to learn from examples and mistakes, to improve continuously with the introduction of new data, and to create a model for individualized patient care [7].

Thanks to the growing informatization of health systems, large amounts of data have become available. The implementation of new technologies and the development of prediction algorithms paved the way for novel possibilities to exploit these huge data collections. Among the several branches of healthcare in which ML aroused enthusiasm, its application in perioperative medicine is showing promising results. In fact, in consideration of its specific characteristics, this analytical technique is suitable for the creation of predictive models, specifically concerning the optimization of resources and the development of warning score systems [8, 9]. The application of these algorithms allows early detection and prediction of acute critical illness, facilitating the management of high-risk patients [10].

More recently, COVID-19 pandemic lighted on the importance of AI-based models for the fast development of algorithms that could integrate readily available data, helping the hospital systems and the clinicians in optimal patient care [11].

The use of ML techniques for the creation of predictive models of perioperative complications is in continuous expansion.

The aim of our review is to clarify the role of ML in perioperative settings, evaluating currently available predictive outcome models, the types of ML algorithms used more frequently, and their proved efficacy.

Methods

Literature search

This systematic review was conducted according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (http://prisma-statement.org/documents/PRISMA_2020_checklist.pdf).

The authors performed a systematic literature search of Scopus, CINAHL, the Cochrane Library, PubMed, MeSH, MEDLINE, and Embase, from 1 January 2015 to 30 March 2021, using different combinations of the following terms: “risk prediction,” “surgery,” “machine learning,” “ICU,” “anesthesia,” and “perioperative.”

Specifically, (((((((("risk prediction"[All Fields]) AND ("surgery"[All Fields])) AND ("machine learning"[All Fields])) OR (risk prediction))) AND (machine learning)) AND (ICU)) OR (risk prediction)) AND (machine learning)) AND (anesthesia); (((((((("risk prediction"[All Fields]) AND ("surgery"[All Fields])) AND ("machine learning"[All Fields])) OR (risk prediction))) AND (machine learning)) AND (ICU)) OR (risk prediction)) AND (machine learning)) AND (anesthesia) OR (((((((("risk prediction"[All Fields]) AND ("surgery"[All Fields])) AND ("machine learning"[All Fields])) OR (risk))) AND (machine learning)) AND (ICU)) OR (risk)) AND (machine learning)) AND (anesthesia); ((postoperative) AND machine learning) AND (intensive care admission).

In the last 10 years, there was an exponential increase in literature concerning the application of AI in medicine. Therefore, we decided to perform the search in this time frame to include more homogeneous and easily comparable studies. We included studies if they evaluated ML models in surgical settings for the prediction of perioperative risk. Both prospective and retrospective studies were eligible for inclusion. The following types of study were excluded: papers published prior to 2015, papers concerning outpatient settings, animal studies, pediatric population, and studies written in languages other than English. Furthermore, primary study evaluating strictly surgical outcomes, and systematic reviews were considered uneligible.

Data extraction and quality assessment

The primary aim of our study was to assess the main perioperative outcomes in which ML methods are used, and their efficacy among different algorithms.

Two reviewers independently screened the selected articles, and a third reviewer resolved any discrepancies.

To assess the reporting quality of all included studies, we used the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) checklist [12]. In fact, it provides guidance for extracting relevant information and calculating summary scores to determine adherence of primary prediction model to the TRIPOD.

Two independent reviewers assessed for each selected study the compliance with the items described in the checklist. Moreover, to facilitate data extraction and scoring, the studies were analyzed according to the study design, predictor selection, outcome assessment, applied model, and its validation. The checklist includes 22 main items, of which ten are divided in sub items, all with four potential answer options: “yes,” “not,” “referenced,” “not applicable.” After adequately fulfilling each item of the checklist, the adherence to the TRIPOD is automatically calculated. We established different levels of adherence to TRIPOD, setting a scale from 0 to 100%, assuming that a research was more accurate with higher adherence to tripod checklist.

Results

One hundred forty-seven papers were identified through database searching. After the removal of the duplicates, 89 articles were screened, and 43 were found to be ineligible after reading the abstracts. Out of the 46 full text reviewed articles, 10 were excluded because of inadequate clinical setting or because concerning pediatric population. Finally, 36 articles were included for the review (Fig. 1).

Outlines all characteristics of the final selected articles (Table 1) [13–48], including the design, cohort, and objective of each study, as well as the ML methods used and the best performance.

Our analyses pointed out that more than 95% of included studies were published after 2018, and almost entirely performed in USA and Asia (Fig. 2).

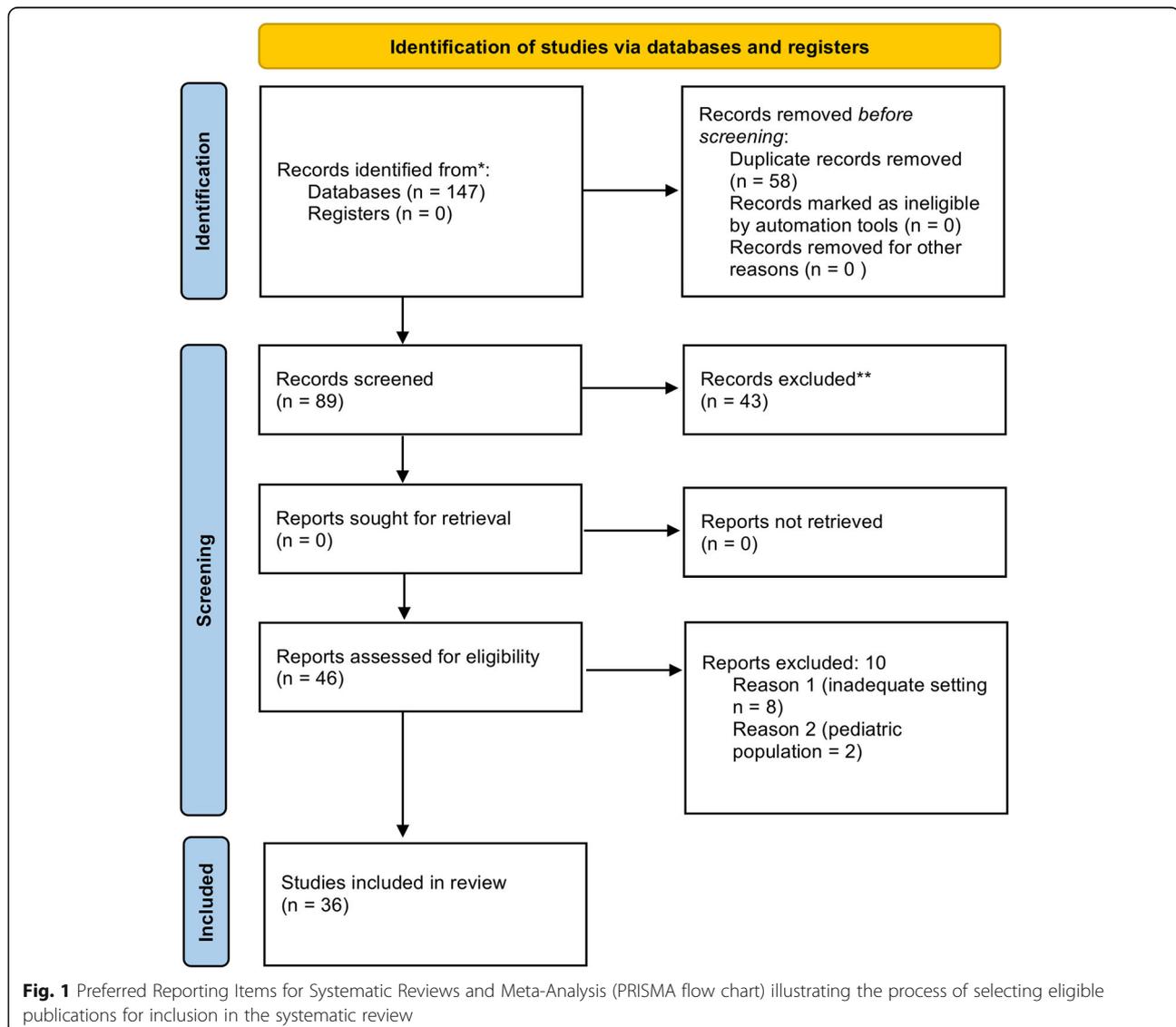


Table 1 Overview of papers included in our analysis

Author, years	Study design	Objective	Final cohort	Outcomes	Type of ML	Prediction performance	Comparator/control
Lundberg SM, 2019	Retrospective/observational single center	Development and testing of a ML-based system that predicts the risk of hypoxemia during general anesthesia	48,069	Hypoxemia	GB	AUC 0.92	ML-based system was compared to anesthesiologists' predictions
Kendale S, 2018 [13]	Retrospective/observational single center	Prediction of the risk of post-induction hypotension using ML methods	13,323	Cardiovascular complications	RF, SVM, GB, BN, LR-EN, regularization, K nearest; linear discrimination analysis; neural nets	AUC GB 0.74 (95% CI, 0.72 to 0.77). RF 0.74 (95% CI, 0.73 to 0.75)	Different ML algorithms were trained to obtain the model with the best performance
Fernandes MPB, 2021 [14]	Retrospective/observational single center	ML models used to predict postoperative mortality rarely include intraoperative factors.	5015	Mortality	logistic regression, RF neural networks, SVM and extreme gradient boosting (XGB).	XGB predicted mortality confidence interval (CI): 0.88 (0.83–0.94)	Different ML algorithms were trained to obtain the model with the best performance
Cherifa M, 2020 [15]	Retrospective/observational single center	Prediction of acute hypotensive episode	1151	Cardiovascular complications	Super Learner (SL) algorithm	SL AUROC 0.890	Different ML algorithms were trained to obtain the model with the best performance
Flechet M, 2019 [16]	Prospective/observational single center	Compare diagnostic performances of ML models and physicians in predicting AKI-23 in the 7 days following ICU admission	252	Acute kidney injury	ML based AKI predictor	AUROC 0.80	Physicians' predictions were compared against the AKI predictor model
Kang AR, 2020 [17]	Retrospective/observational single center	Prediction of hypotension during anesthesia induction	222	Cardiovascular complications	Naïve Bayes, logistic regression, RF, ANN	RF best performance AUC 0.842	Different ML algorithms were trained to obtain the model with the best performance
Meiring C, 2018 [18]	Retrospective/observational multicentric	Identification of risk factors for admission in ER/ICU for spine patients	11150	ER/ ICU admission	RF, SVM, GB, DECISION TREE, DEEP LEARNING, NNC, Single layer averaged neural network	RF AUC 0.859, NNC AUC 0.864; SVM AUC 0.867; adaboost AUC 0.868; deep learning AUC 0.883	Logistic regression against physiological data alone outperformed APACHE-II (current risk stratification tools)
Nudel J, 2021 [17]	Retrospective/observational multicentric	Comparison of two ML strategies with conventional statistical models in prediction of surgical complication	43,6807	Surgical complications, VTE	GB, ANN	ANN, and XGB, LR achieved similar AUCs 0.65, 0.67 and 0.64	Different ML algorithms were trained to obtain the model with the best performance
Lee Hc, 2018 [19]	Retrospective/observational single center	Comparison of ML method with logistic regression analysis to predict AKI after cardiac surgery	2010	AKI, mortality	RF, SVM, GB, DECISION TREE, DEEP LEARNING, NNC	Best GB AUC 0.78	The performance of ML approaches was compared with that of LR analysis
Bai P, 2020 [20]	Retrospective/observational multicentric	Identification of risk factors of early cerebral infarction and myocardial infarction after CEA with ML method	443	Cardiovascular complications	linear SVM, decision tree, RF, ANN, quadratic discriminant analysis, and XGBoost	XGBoost had the highest accuracy	Not applicable
Tan HS, 2021 [2021]	Retrospective study single center	Use of ML to identify predictive factors for inadequate labor anesthesia	20,716	Pain prevention	RF, XGBoost and logistic regression models	All three models performed similarly, with AUC 0.763–0.772	The performance of ML was compared with regression techniques
Solomon	Retrospective	Prediction of	62,182	Cardiovascular	Gradient	AUC of 0.81–	The performance of ML

Table 1 Overview of papers included in our analysis (Continued)

Author, years	Study design	Objective	Final cohort	Outcomes	Type of ML	Prediction performance	Comparator/control
SC, 2020 [21]	and prognostic single center	intraoperative bradycardia		complications	Boosting Machine (GBM)	0,89	was compared with regression techniques
Ko S, 2020 [22]	Retrospective and multicentric	Prediction of postoperative AKI after total knee arthroplasty.	5757	AKI	Gradient Boosting Machine (GBM)	AUC of 0,78	Not applicable
Lu Y, 2020 [23]	Retrospective single center	Develop ML algorithm for identification of patients requiring admission following elective anterior cruciate ligament (ACL) reconstruction.	4709	Length of stay	RF, XGBoost, LDA, AdaBoost	The ensemble model achieved the best AUC 0.76	Not applicable
Maheshwari K, 2020 [24]	Observational single center	Using ML to predict intraoperative hypotension	305	Cardiovascular complications	Hypotension Prediction Index	95% confidence interval	Not applicable
Hill BL, 2019 [25]	Retrospective/observational single center	Develop a model that estimates in-hospital mortality at the end of surgery to quantify the change in risk during the perioperative period.	53,097	Mortality	Logistic regression, Elastic Net24 logistic regression, RF, GB.	Best RF 0.932	Comparison of ML methods with the perioperative score (as ASA physical status score)
Suhre W, 2020 [26]	Retrospective multicentric	Correlation between chronic cannabis use and the risk of postoperative nausea and vomiting (PONV).	16,245	PONV	Bayesian additive regression trees (BART)	90% CI 0.98–1.33	Not applicable
Lee HC, 2018 [27]	Retrospective/observational single center	Comparison of ML method with logistic regression analysis to predict AKI after liver transplantation	1211	AKI, mortality	RF, SVM, GB, Decision tree, Neural network Classifier, BN, LR-EN, multilayer perceptron	Best GB AUC 0.90	The performance of ML approaches was compared with that of LR analysis
Barry GS, 2021 [28]	Retrospective cohort study	Investigate the incidence and factors associated with rebound pain in patients who received a PNB for ambulatory surgery.	482	Pain control	Logistic model tree attribute-selected classifier	ROC curve of 0.609	Not applicable
Gabriel RA, 2019 [29]	Retrospective/observational single center	Develop a predictive model for determining LOS.	1018	LOS	Ridge regression, Lasso, RF	ridge regression 0.761, Lasso 0.752, RF 0.731	Predictive models using ML techniques were compared to model performances
Li H, 2020 [30]	Retrospective/observational single center	Development of a predictive model for LOS after total knee arthroplasty	1826	LOS	GB	AUC 0.738.	Logistic regression and ML model were compared
Jungquist CR, 2019 [31]	Retrospective/observational single center	Early detection of respiratory depression using ML models	60	Postoperative respiratory complications	SVM	Accuracy of 80%	Not applicable
Nguyen M, 2020 [32]	Multicentric randomized	Using ML techniques and causal inference methods to detect the dynamic relationship between transfusion ratios and outcomes in trauma patients	680	Mortality and hemorrhagic complications	Statistical programming language R	Mortality at AUC 0.89, hemorrhagic complications 1.07	ML techniques were used to augment the intent-to-treat analysis of PROPPR
Tourani R, 2019 [33]	Retrospective multicentric	In the context of perioperative decision support, understand if the use of intraoperative data improve the performance	38,045 + 9,044	Sepsis, septic shock, UTI, PNA, surgical infections	Logistic regression models.	AUC between 0.66 and 0.82	Not applicable

Table 1 Overview of papers included in our analysis (Continued)

Author, years	Study design	Objective	Final cohort	Outcomes	Type of ML	Prediction performance	Comparator/control
Cartailler J, 2019 [34]	Clinical trial single center	Use of EEG-patterns to anticipate excessive deep sedation	80	Neurological complications	RF	AUC of 0.93	Not applicable
Wong WEJ, 2021 [35]	Retrospective/observational single center	Prediction of AKI in ICU	940	ICU AKI, hospital and 1 year mortality	Chi-square test, Fisher's exact test, Mann-Whitney test, independent <i>t</i> test and the Kaplan-Meier curve.	AUROC of the auxiliary models for ICU AKI were 0.7537, 0.7589, 0.7950, 0.7333 and 0.7654.	Not applicable
Lee CK, 2021 [36]	Retrospective/observational single center	Prediction of mortality in post-operative patients	59,985	Post-operative mortality	Generalized additive models with neural networks (GAM-NNs).	AUC 0.921	Model performance was compared to a standard LR model
Jeong YS, 2021 [37]	Retrospective/observational single center	To make a proper model for predicting postoperative major cardiac event (MACE) in ESRD patients undergoing general anesthesia.	3220	Cardiovascular complications, mortality	SVM, decision tree, RF, Gaussian naive Bayes (GNB), ANN, LR, XGBoost	RF AUC 0.797	Different ML algorithms were trained to obtain the model with the best performance
Filiberto AC, 2021 [38]	Retrospective/observational single center	Postoperative acute kidney injury using ML models	1531	AKI	RF	AUC 0.70	ML models using the perioperative data were compared to models using either preoperative data alone or the ASA physical status classification
Meyer A, 2018 [39]	Retrospective/observational single center	Use machine learning methods to predict severe complications during and after cardiothoracic surgery.	11,492	Postoperative bleeding, AKI, mortality	Deep learning model	AUC 0.09 for bleeding, of 0.18 for mortality, and of 0.25 for AKI	Deep learning methods were compare to established standard-of-care clinical reference tools
Chiew CJ, 2020 [40]	Retrospective/observational single center	Compare the performance of ML models against the traditionally (CARES) model and (ASA-PS) in the prediction of 30-day postsurgical mortality and ICU admission	90,785	Mortality, postoperative ICU admission	RF, GB, adaptive boosting, SVM	Best GB AUC 0.23 and for mortality and 0.38 ICU admission	The performance of ML models was compare against the traditionally Combined Assessment of Risk and Encountered in Surgery (CARES) model and the ASA physical status.
Bihorac A, 2019 [41]	Retrospective/observational single center	To calculate the risk for postoperative complications and death after surgery using ML	51,457	AKI, sepsis, VTE, ICU admission > 48 h, mechanical ventilation > 48 h, wound, neurologic and cardiovascular complications	MySurgeryRisk algorithm	AUC values ranging between 0.82 and 0.94	Not applicable
Yao RQ, 2020 [42]	Retrospective/observational single center	Develop a mathematical model for predicting the in-hospital mortality among patients with postoperative sepsis.	3713	Postoperative sepsi, mortality	Extreme gradient boosting (XGBoost) and stepwise logistic regression	Best XGBoost AUC 0.835	ML model was compare to the stepwise LR model.
Datta S, 2020 [43]	Retrospective/observational single center	Describe a model that predicts postoperative complications considering	43,943	ICU LOS, prolonged mechanical	RF	AUC 0.21	ML models using preoperative and intraoperative data were

Table 1 Overview of papers included in our analysis (Continued)

Author, years	Study design	Objective	Final cohort	Outcomes	Type of ML	Prediction performance	Comparator/control
		intraoperative events.		ventilation, neurologic complications cardiovascular complications, AKI, VTE, wound complications, mortality			compare to models using preoperative data alone
Brennan M, 2019 [44]	Prospective, non-randomized pilot study	Assess the usability and accuracy of the MySurgeryRisk algorithm for preoperative risk assessment	20	AKI, sepsis, VTE, ICU admission > 48 h, mechanical ventilation > 48 h, wound, neurologic and cardiovascular complications	MySurgeryRisk algorithm	MySurgeryRisk algorithm ranged between 0.73 and 0.85	Compare the accuracy of perioperative risk-assessment between physicians and MySurgeryRisk.
Houthoof R, 2015 [45]	Retrospective/observational single center	develop model to determine patient survival and ICU length of stay (LOS) based on monitored ICU patient data.	14,480	LOS	ANN, k-nearest neighbors (k-NN), SVMs, classification trees (CART), RF, adaptive boosting (AdaBoost)	SVM AUC 0.77	Different ML algorithms were trained to obtain the model with the best performance

AdaBoost = adaptive boosting algorithms, *AKI* = acute kidney injury; *ANN* = artificial neural network models, *BART* = Bayesian additive regression trees, *BN* = Bayesian network, *GB* = gradient boosting, *ICU* = intensive care unit, *LDA* = linear discriminant classifier, *LOS* = length of stay, *LR-EN* = logistic regression with elastic net, *ML* = machine learning, *NNC* = neural network classifier, *PNA* = pneumonia, *PONV* = postoperative nausea and vomiting, *RF* = Random Forest, *SVM* = support vector machine, *UTI* = urinary tract infection, *VTE* = venous thromboembolism, *XGBoost* = extreme gradient boosting, *ASA* = American Society of Anesthesiologist

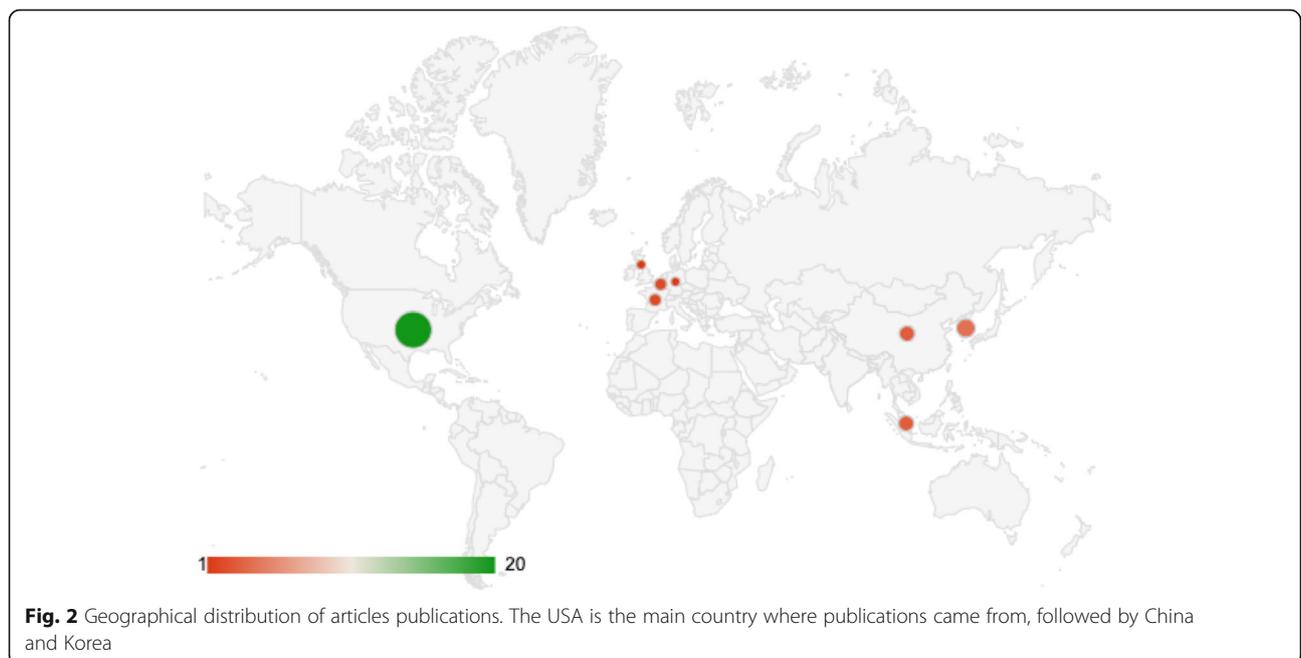
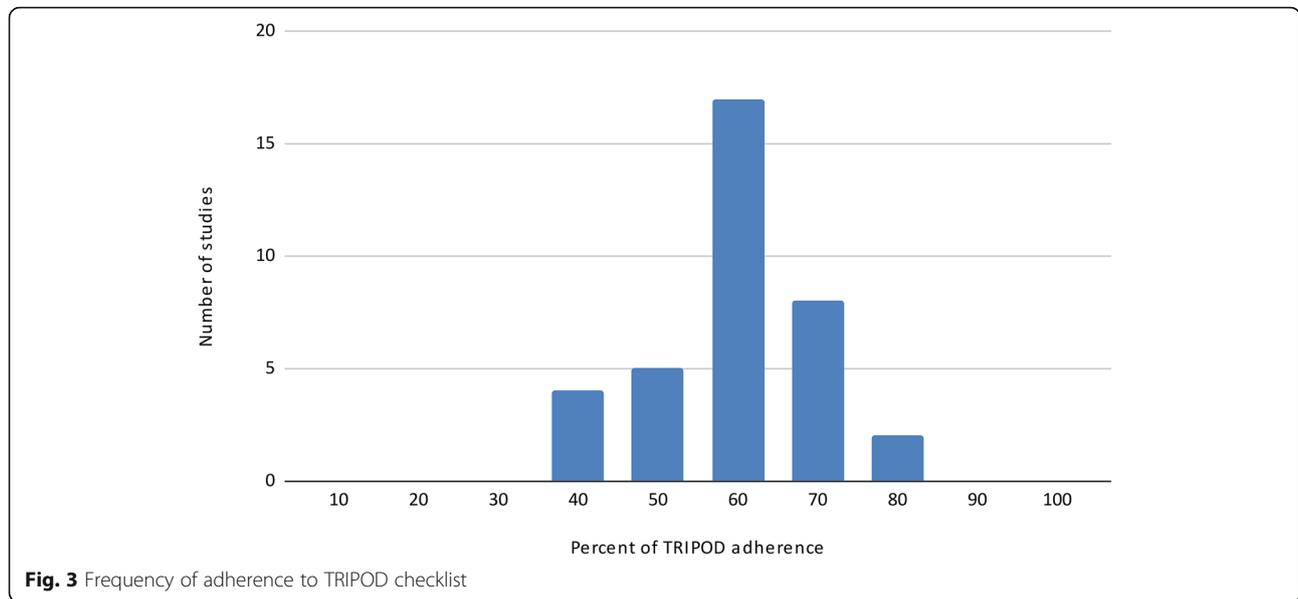


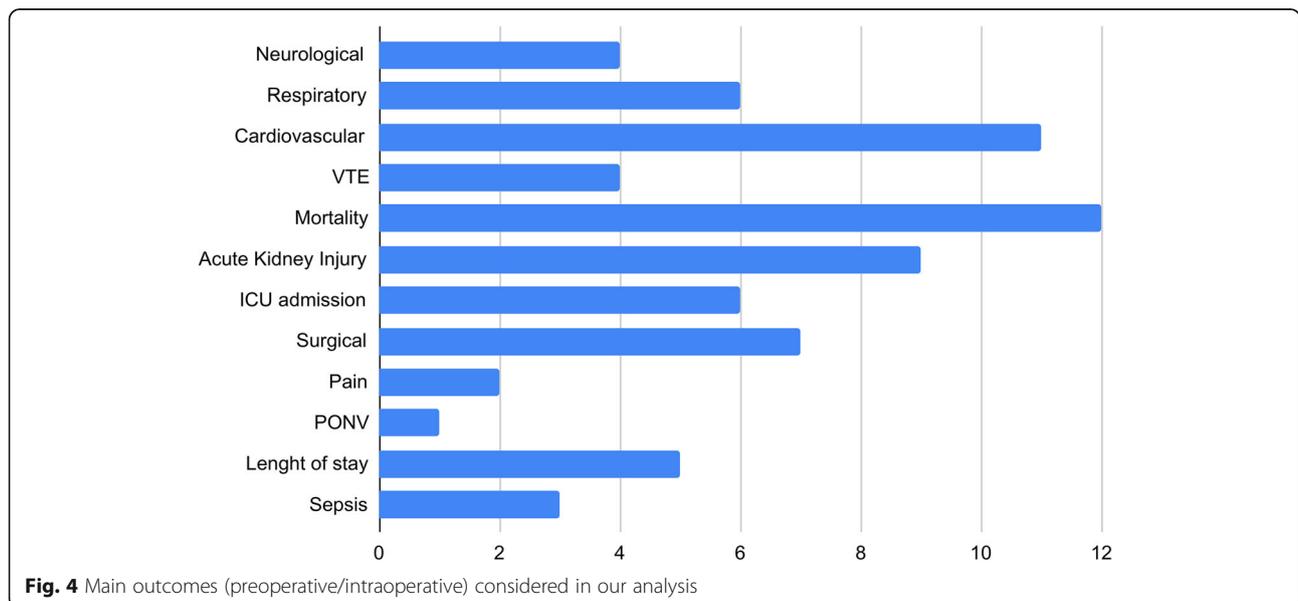
Fig. 2 Geographical distribution of articles publications. The USA is the main country where publications came from, followed by China and Korea



The quality of the studies selected for the review was acceptable, with 75% of studied showing an adherence rate to TRIPOD more than 60% (Fig. 3). Specifically, in the first section of the checklist (Title and Abstract), a mean of 42% of studies adhere to tripod item. Concerning the methods section, all the articles defined the study design, or the source of data, while 53% of papers described the handling of missing data. In the results section, measures applied and models used were not always appropriated in the included studies, specifically 8% of papers presented the full prediction model and explained how to use it, while 19% of studies reported performance measures for the prediction model (Rev #2, comm #3).

Nearly all manuscripts discussed about the limitations of the study and gave an overall interpretation of results.

The use of these new technologies to analyze perioperative complications has been tested in almost all types of surgery (general, cardiac, orthopedic, neurosurgical, vascular). Variables and predictors were properly listed and described in all the articles. ML methods were used mainly to predict the following outcomes: mortality ($n = 12$), cardiovascular complications ($n = 11$), acute kidney injury (AKI; $n = 9$), surgical complications ($n = 7$), intensive care unit admission (ICU; $n = 6$), respiratory complications ($n = 6$), length of stay ($n = 5$), venous thromboembolism (VTE; $n = 4$), neurological



complications ($n = 4$), sepsis ($n = 3$), pain ($n = 2$), and post-operative nausea and vomiting (PONV; $n = 1$) (Fig. 4).

As stated before, most of studies considered preoperative variables, like demographic, medical history, clinical and laboratory values evaluation, to calculate perioperative risk. Conversely, several studies evaluated intraoperative variables, as electroencephalography (EEG) pattern [34], or intraoperative vital signs [13, 15, 22, 24, 46, 47], for a real-time prediction of overly deep sedation, post-induction and intraoperative hypotension, hypoxemia, and intraoperative bradycardia.

Supervised models were used in most of cases (Fig. 5). The most frequently used algorithms were gradient boosting ($n = 13$), random forest ($n = 10$), logistic regression (LR; $n = 7$), artificial neural networks (ANNs; $n = 6$), and support vector machines (SVM; $n = 6$). Deep learning, decision trees, and Naïve Bayes were other models commonly applied in the included manuscripts.

In the totality of reviewed papers, ML algorithms proved to be effective in outcome prediction. Half of the selected studies compared different types of ML to identify the best performing method. Gradient boosting and random forest were found to be the models with the highest accuracy, achieving an area under the curve (AUC) greater than 0.90 in most of cases. Moreover, a few studies compared automatically obtained algorithms to conventional scores, revealing the outperformance of ML models [25].

Discussion

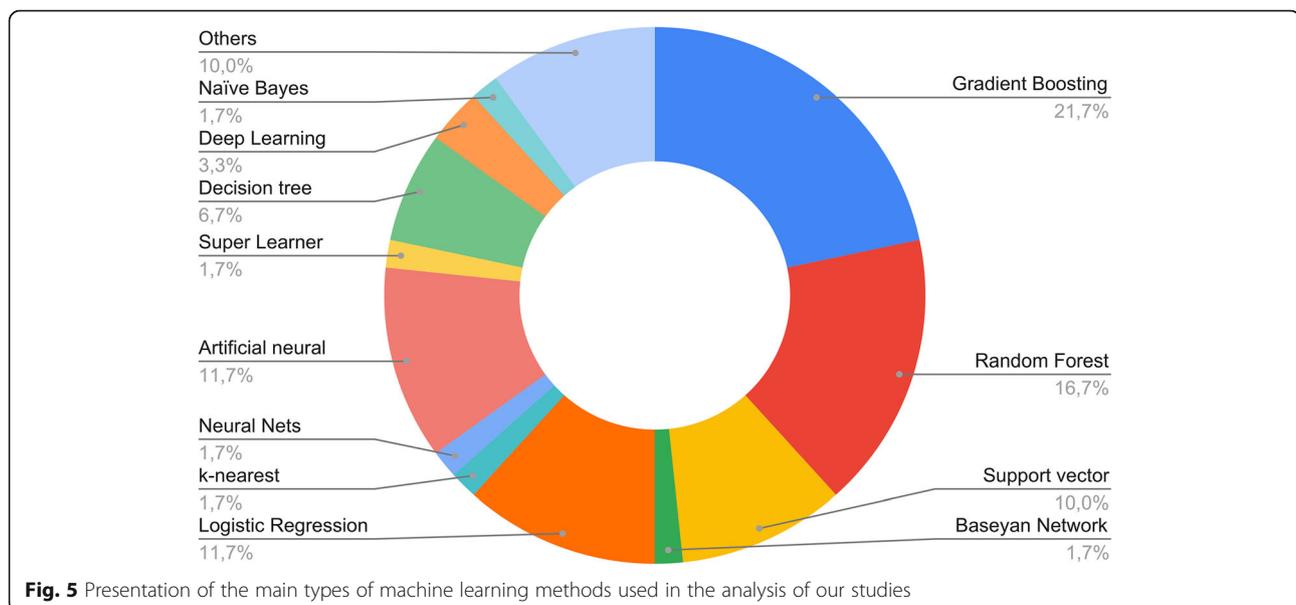
The number of manuscripts regarding ML implementation in health care settings is steadily increasing over the last few years, as clearly suggested by a recently

published review on AI utility to provide decision support to clinicians in ICU setting [49, 50].

In fact, the availability of electronic health records, and the diffusion of Big Data systems have enabled new possibilities in data collection and storage. The interpretation of this amount of data with traditional methods could not only be extremely complicated, but even reductive. In this regard, the advent of AI-based technologies has opened up new perspectives, providing a different form of research [51].

Anesthesia and assessment of perioperative risk appear to be excellent fields to develop and apply ML systems, as reported in literature [52, 53], and confirmed by our research. The identification of modifiable risk factors and the subsequent optimization of the preoperative phase appear to be a crucial factor to decrease the incidence of post-operative complications [54]. Furthermore, risk stratification allows the acquisition of an adequate informed consent and an accurate anesthesiologic planning, tailored to each patient. ML systems are well suitable for this context, where the possibility to collect a large number of data and the choice of the variable that is selected by the model itself, allows the discovery of new factors and a different interpretation of already known items. Thus, the availability of interpretations and predictions in real time could allow to enter a new era of anesthesia.

From a practical point of view, the method starts with multi-source data extrapolated and collected; subsequently, they are placed in ML systems able to return interpretative and predictive models, providing suitable tools for daily technologies with validated scores. Among conventional scores, the one used more frequently for comparison is the ASA-PS Classification System that has



been in use for over 60 years. Comparing existing scores with new models is an essential step to understand whether this investment of time and resources could finally improve the perioperative risk stratification.

Moreover, in addition to the risk of post-operative complications, ML would also be able to answer more complex questions and create models capable of providing early predictions of adverse events, thus enabling a perioperative optimization.

The results that emerge from this systematic analysis are promising. In studies that compared ML models with traditional scores, most confirmed their outperformance. In particular, the use of AI-based technologies provided excellent results regarding events of great interest in the field of Anesthesia, as post-induction hypotension and post-intubation hypoxia [13], or the risk of AKI or delirium after surgery [19, 27, 55].

Finally, it is interesting to underline that not only clinical outcomes are relevant, but also administrative ones, as length of hospital stay, or need for recovery in intensive care settings, that may have a great relapse into hospital logistics and in economic strategies (Fig. 6). A systematic use of AI might allow the achievement of

innovative results in other fields as well, such as scientific research and health organization, especially when associated with other data management technologies such as Big Data and Blockchain.

Among several ML algorithms currently applied, Gradient boosting and random forest were found to be the models with the best performance and the highest accuracy, achieving an area under the curve (AUC) greater than 0.90 (Ref #2, comm #3). Still, it is not possible to make a uniform evaluation and draw conclusions about the best algorithm for predictive models of perioperative complications, because of the heterogeneity of settings and the difference in the algorithms evaluated. The lack of uniformity of the included studies prevented us from performing a meta-analysis using univariate and multivariate random effect models (Ref #2, comm #3). Moreover, the models in most of the studies lack an external validation.

Further, even if we practically use AUC as an evaluation criterion, we acknowledge its limits in the setting of AI, especially in case of unbalanced dataset. Note that other criteria can also be used to evaluate ML models, such as model relevance, efficiency, and interpretability

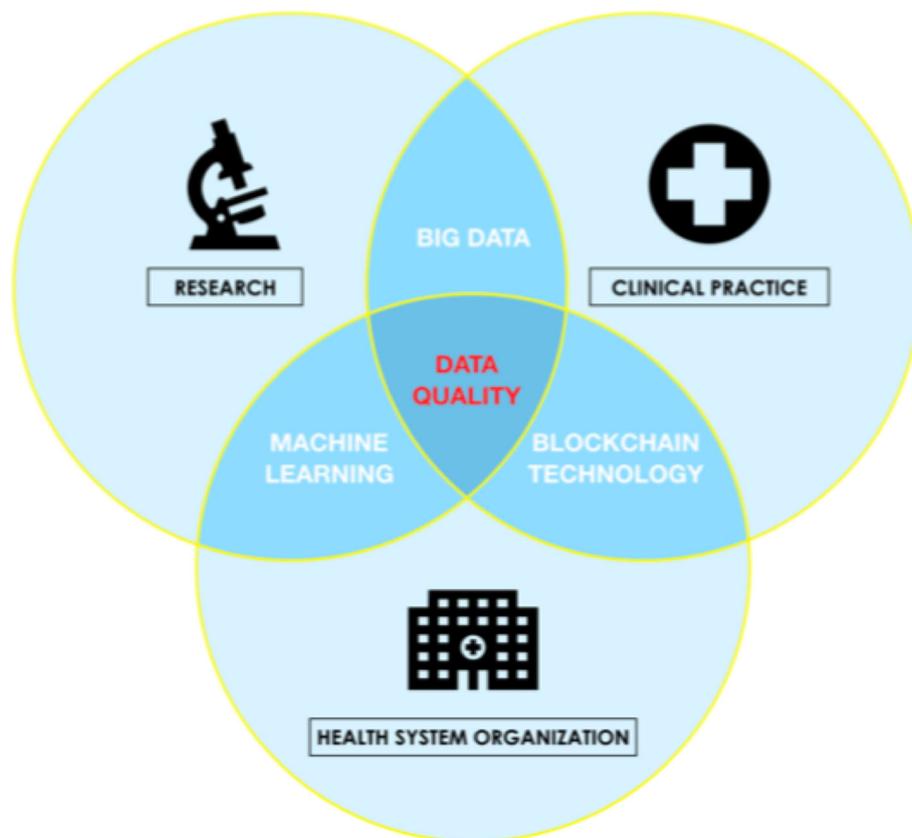


Fig. 6 Importance of acquisition of data quality for application of AI in different fields such as research, clinical practice, and health system organization

[56]. However, to achieve high-quality and high-quantity data sets, it is of paramount importance the screening of each step of the process, from data collection to ML model selection and its algorithm

(Rev #2, comm #3, comm #4).

Despite their growing diffusion, the use of these technologies in perioperative medicine is raising limitations and challenges. Along with technological progress, data quality will inevitably become increasingly important. A viable choice could be blockchain technology, to ensure adequate quality and enable secure data sharing. Its implementation could allow the safe management of large files and consequently the approval of algorithms that are progressively developed [57].

Furthermore, as recently reported for ICU-setting [50], despite the potential role of AI to improve clinical outcomes, the vast majority of developed models remain within the testing and prototyping environment. A uniform and structured approach could enable the implementation and safe delivery of AI technologies in ICU and overall, in health care settings.

Finally, the creation of predictive scores should guarantee precise rules. Unfortunately, these technologies are so innovative that the evaluation of their performance is not always so linear. Therefore, a new version of the TRIPOD statement specific for AI/ML systems (TRIPOD-ML) is currently under development. It will focus on the introduction of ML prediction algorithms to establish methodological and reporting standards for ML studies in health care [58].

Technologies are becoming more and more present in health-care settings. Both clinical and organizational decision-making processes can take advantage of these technologies. Nevertheless, high-quality studies are needed to demonstrate the real impact of ML in this context.

Our research group is starting a study that aims to validate a safe discharge score from the PACU (post-anesthesia care unit) using AI techniques; the score will no longer be generic, but based on the local clinical reality and on the specific population. Similarly, we are working on the application of AI algorithms in OR (operating room) management settings, developing a prospective trial “Bloc-op” (NCT 05106621), in collaboration with the engineering department, to optimize OR organization and resources allocation. We believe that multidisciplinary collaboration is essential to integrate AI technologies into routine clinical practice, thus leading to a great improvement in the quality of care.

We proposed that AI should become an essential technical and non-technical skill for the future anesthesiologists. In order to achieve this goal, a primary focus should be the education and training of physicians and

researchers, who need to be adequately prepared on the uses and limitations of AI based technologies (Rev #2, comm #4).

Conclusions

This systematic review shows the potential role of ML in perioperative medicine, and particularly in the creation of models for the prediction of perioperative risk. Our results are encouraging.

Undoubtedly, the exploitation of a large amount of data is possible solely thanks to the application of AI. ML algorithms offer increasingly precise solutions in terms of optimization of the perioperative risk. A personalized risk/benefit analysis can result in an accurate prediction in terms of length of hospital stay and ICU recovery, thus positively influencing patient management and health costs.

Further research is needed to develop a framework standardizing AI evaluation measures, and this will be possible with interdisciplinary approaches, allowing to constantly improve high-quality care.

Abbreviations

ML: Machine learning; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses; ICU: Intensive care unit; TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis; AKI: Acute kidney injury; LR: Logistic regression; ANNs: Artificial neural networks; SVM: Support vector machines; ASA-PS: American Society of Anaesthesiologists Physical Status; EuroSCORE: European system for cardiac operative risk evaluation; AI: Artificial intelligence; EEG: Electroencephalography; AUC: Area under the curve; TRIPOD-ML: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis-Machine Learning; PACU: Post-anesthesia care unit; OR: Operating room

Acknowledgements

Not applicable

Authors' contributions

VB, MV, PDR, and EB selected and identified the eligible studies. MM, GL, and EB analyzed and interpreted the data. VB and EB identified the aim of the review and revised the work. GB, BP, and MC elaborated, drafted, and revised the work. All authors read and approved the final manuscript. All authors have agreed to be responsible for the content of the work, and to ensure that questions related to the accuracy or integrity of any part of work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Funding

This review was completed as a part of the research fellowship of University of Parma MADA-MED (MACHINE learning and big DATA in medicina perioperatoria), co-funded with resources of FSE (Fondo Sociale Europeo, delibera di G.R. 589/2019-Rif. PA 2019-11449/RER). This research has financially been supported by the Programme “FIL-Quota Incentivante” of University of Parma and co-sponsored by Fondazione Cariparma.

Availability of data and materials

Not applicable

Declarations

Competing interest

All authors declare that they have no competing interests.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Author details

¹Anesthesiology, Critical Care and Pain Medicine Division, Department of Medicine and Surgery, University of Parma, Viale Gramsci 14, 43126 Parma, Italy. ²General Surgery Unit, Department of Medicine and Surgery, University of Parma, Viale Gramsci 14, 43126 Parma, Italy. ³Department of Engineering and Architecture, University of Parma, Viale G.P.Usberti 181/A, 43124 Parma, Italy.

Received: 7 December 2021 Accepted: 4 January 2022

Published online: 15 January 2022

References

- Bose S, Talmor D. (2018) Who is a high-risk surgical patient? *Curr Opin Crit Care*;24(6):547-553. doi: <https://doi.org/10.1097/MCC.0000000000000556>. PMID: 30308542.
- Rajkomar A, Dean J, Kohane I. (2019) Machine Learning in Medicine. *N Engl J Med*. 380(14):1347-1358. doi: <https://doi.org/10.1056/NEJMr1814259>. PMID: 30943338.
- Wiijesundera DN (2015) Predicting outcomes: Is there utility in risk scores? *Can J Anaesth* 2016 63(2):148-158. <https://doi.org/10.1007/s12630-015-0537-2> Epub PMID: 26670801
- Mayhew D, Mendonca V, Murthy BVS (2019) A review of ASA physical status - historical perspectives and modern developments. *Anaesthesia* 74(3):373-379. <https://doi.org/10.1111/anae.14569> Epub 2019 Jan 15. PMID: 30648259
- Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. (1999). European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg*. 16(1):9-13. doi: [https://doi.org/10.1016/s1010-7940\(99\)00134-7](https://doi.org/10.1016/s1010-7940(99)00134-7). PMID: 10456395.
- Kheterpal S, Tremper KK, Heung M, Rosenberg AL, Englesbe M, Shanks AM, Campbell DA Jr. (2009). Development and validation of an acute kidney injury risk index for patients undergoing general surgery: results from a national data set. *Anesthesiology*. 110(3):505-515. doi: <https://doi.org/10.1097/ALN.0b013e3181979440>. PMID: 19212261.
- Beam AL, Kohane IS. (2017). Big Data and Machine Learning in Health Care. *JAMA* ;319(13):1317-1318. doi: <https://doi.org/10.1001/jama.2017.18391>. PMID: 29532063.
- Lauritsen SM, Kristensen M, Olsen MV, Larsen MS, Lauritsen KM, Jørgensen MJ, Lange J, Thiesson B. (2020). Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nat Commun*; 11(1):3852. doi: <https://doi.org/10.1038/s41467-020-17431-x>. PMID: 32737308; PMCID: PMC7395744.
- Xue B, Li D, Lu C, King CR, Wildes T, Avidan MS, Kannampallil T, Abraham J. (2021). Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. *JAMA Netw Open*;4(3):e212240. doi: <https://doi.org/10.1001/jamanetworkopen.2021.2240>. PMID: 33783520; PMCID: PMC8010590.
- Hyland SL, Faltsy M, Hüser M, Lyu X, Gumbsch T, Esteban C, Bock C, Horn M, Moor M, Rieck B, Zimmermann M, Bodenham D, Borgwardt K, Rättsch G, Merz TM (2020) Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med* 26(3):364-373. <https://doi.org/10.1038/s41591-020-0789-4> Epub 2020 Mar 9. PMID: 32152583
- Alimadadi A, Aryal S, Manandhar I, Munroe PB, Joe B, Cheng X (2020) Artificial intelligence and machine learning to fight COVID-19. *Physiol Genomics* 52(4):200-202. <https://doi.org/10.1152/physiolgenomics.00029.2020> Epub. PMID: 32216577; PMCID: PMC7191426
- Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*;162(1):W1-73. doi: <https://doi.org/10.7326/M14-0698>. PMID: 25560730.
- Kendale S, Kulkarni P, Rosenberg AD, Wang J. (2018). Supervised Machine-learning Predictive Analytics for Prediction of Postinduction Hypotension. *Anesthesiology*;129(4):675-688. doi: <https://doi.org/10.1097/ALN.0000000000002374>. PMID: 30074930.
- Fernandes MPB, Armengol de la Hoz M, Rangasamy V, Subramaniam B (2021) Machine Learning Models with Preoperative Risk Factors and Intraoperative Hypotension Parameters Predict Mortality After Cardiac Surgery. *J Cardiothorac Vasc Anesth* 35(3):857-865. <https://doi.org/10.1053/j.jvca.2020.07.029> Epub 2020 Jul 12. PMID: 32747203
- Cherifa M, Blet A, Chambaz A, Gayat E, Resche-Rigon M, Pirracchio R. (2020). Prediction of an acute hypotensive episode during an ICU hospitalization with a super learner machine-learning algorithm. *Anesth Analg*; 130(5):1157-1166. doi: <https://doi.org/10.1213/ANE.0000000000004539>. PMID: 32287123.
- Flechet M, Falini S, Bonetti C, Güiza F, Schetz M, Van den Berghe G, Meyfroidt G (2019) Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Crit Care* 23(1):282. <https://doi.org/10.1186/s13054-019-2563-x> PMID: 31420056; PMCID: PMC6697946
- Nudel J, Bishara AM, de Geus SWL, Patil P, Srinivasan J, Hess DT, Woodson J (2021) Development and validation of machine learning models to predict gastrointestinal leak and venous thromboembolism after weight loss surgery: an analysis of the MBSAQIP database. *Surg Endosc* 35(1):182-191. <https://doi.org/10.1007/s00464-020-07378-x> Epub 2020 Jan 17. PMID: 31953733
- Meiring C, Dixit A, Harris S, MacCallum NS, Brealey DA, Watkinson PJ, Jones A, Ashworth S, Beale R, Brett SJ, Singer M, Ercole A (2018) Optimal intensive care outcome prediction over time using machine learning. *PLoS One* 13(11):e0206862. <https://doi.org/10.1371/journal.pone.0206862> PMID: 30427913; PMCID: PMC6241126
- Lee HC, Yoon HK, Nam K, Cho YJ, Kim TK, Kim WH, Bahk JH (2018) Derivation and validation of machine learning approaches to predict acute kidney injury after cardiac surgery. *J Clin Med* 7(10):322. <https://doi.org/10.3390/jcm7100322> PMID: 30282956; PMCID: PMC6210196
- Bai P, Zhou Y, Liu Y, Li G, Li Z, Wang T, Guo X (2020, 2020) Risk Factors of Cerebral Infarction and Myocardial Infarction after Carotid Endarterectomy Analyzed by Machine Learning. *Comput Math Methods Med*:6217392. <https://doi.org/10.1155/2020/6217392> PMID: 33273961; PMCID: PMC7683166
- Solomon SC, Saxena RC, Neradilek MB, Hau V, Fong CT, Lang JD, Posner KL, Nair BG. (2020) Forecasting a Crisis: Machine-Learning Models Predict Occurrence of Intraoperative Bradycardia Associated With Hypotension. *Anesth Analg*;130(5):1201-1210. doi: <https://doi.org/10.1213/ANE.0000000000004636>. PMID: 32287127.
- Ko S, Jo C, Chang CB, Lee YS, Moon YW, Youm JW, Han HS, Lee MC, Lee H, Ro DH (2020) A web-based machine-learning algorithm predicting postoperative acute kidney injury after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc*. <https://doi.org/10.1007/s00167-020-06258-0> Epub ahead of print. PMID: 32880677
- Lu Y, Forlenza E, Cohn MR, Lavoie-Gagne O, Wilbur RR, Song BM, Krych AJ, Forsythe B. (2020). Machine learning can reliably identify patients at risk of overnight hospital admission following anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc* 2021;29(9):2958-2966. doi: <https://doi.org/10.1007/s00167-020-06321-w>. Epub. PMID: 33047150.
- Maheshwari K, Buddi S, Jian Z, Settels J, Shimada T, Cohen B, Sessler DI, Hatib F (2021) Performance of the Hypotension Prediction Index with non-invasive arterial pressure waveforms in non-cardiac surgical patients. *J Clin Monit Comput* 35(1):71-78. <https://doi.org/10.1007/s10877-020-00463-5> Epub 2020 Jan 27. PMID: 31989416; PMCID: PMC7889685
- Hill BL, Brown R, Gabel E, Rakocz N, Lee C, Cannesson M, Baldi P, Olde Loohuis L, Johnson R, Jew B, Maoz U, Mahajan A, Sankaraman S, Hofer I, Halperin E (2019) An automated machine learning-based model predicts postoperative mortality using readily-extractable preoperative electronic health record data. *Br J Anaesth* 123(6):877-886. <https://doi.org/10.1016/j.bja.2019.07.030> Epub. PMID: 31627890; PMCID: PMC6883494
- Suhre W, O'Reilly-Shah V, Van Cleve W (2020) Cannabis use is associated with a small increase in the risk of postoperative nausea and vomiting: a retrospective machine-learning causal analysis. *BMC Anesthesiol* 20(1):115. <https://doi.org/10.1186/s12871-020-01036-4> PMID: 32423445; PMCID: PMC7236204
- Lee HC, Yoon SB, Yang SM, Kim WH, Ryu HG, Jung CW, Suh KS, Lee KH (2018) Prediction of acute kidney injury after liver transplantation: machine learning approaches vs. logistic regression model. *J Clin Med* 7(11):428. <https://doi.org/10.3390/jcm7110428> PMID: 30413107; PMCID: PMC6262324
- Barry GS, Bailey JG, Sardinha J, Brousseau P, Uppal V (2021) Factors associated with rebound pain after peripheral nerve block for ambulatory surgery. *Br J Anaesth* 126(4):862-871. <https://doi.org/10.1016/j.bja.2020.10.035> Epub 2020 Dec 31. PMID: 33390261

29. Gabriel RA, Sharma BS, Doan CN, Jiang X, Schmidt UH, Vaida F. (2019). A predictive model for determining patients not requiring prolonged hospital length of stay after elective primary total hip arthroplasty. *Anesth Analg*; 129(1):43–50. doi: <https://doi.org/10.1213/ANE.0000000000003798>. PMID: 30234533.
30. Li H, Jiao J, Zhang S, Tang H, Qu X, Yue B (2020) Construction and comparison of predictive models for length of stay after total knee arthroplasty: regression model and machine learning analysis based on 1,826 cases in a single Singapore center. *J Knee Surg*. <https://doi.org/10.1055/s-0040-1710573> Epub ahead of print. PMID: 32512596
31. Jungquist CR, Chandola V, Spulecki C, Nguyen KV, Crescenzi P, Tekeste D, Sayapaneni PR (2019) Identifying patients experiencing opioid-induced respiratory depression during recovery from anesthesia: the application of electronic monitoring devices. *Worldviews Evid Based Nurs* 16(3):186–194. <https://doi.org/10.1111/wvn.12362> Epub 2019 May 2. PMID: 31050151
32. Nguyen M, Pirracchio R, Kornblith LZ, Callcut R, Fox EE, Wade CE, Schreiber M, Holcomb JB, Coyle J, Cohen M, Hubbard A (2020) Dynamic impact of transfusion ratios on outcomes in severely injured patients: targeted machine learning analysis of the Pragmatic, Randomized Optimal Platelet and Plasma Ratios randomized clinical trial. *J Trauma Acute Care Surg* 89(3): 505–513. <https://doi.org/10.1097/TA.0000000000002819> PMID: 32520897; PMID: PMC7830749
33. Tourani R, Murphree DH, Melton-Meaux G, Wick E, Kor DJ, Simon GJ (2019) The value of aggregated high-resolution intraoperative data for predicting post-surgical infectious complications at two independent sites. *Stud Health Technol Inform* 264:398–402. <https://doi.org/10.3233/SHIT190251> PMID: 31437953; PMID: PMC7037580
34. Cartailier J, Parutto P, Touchard C, Vallée F, Holcman D (2019) Alpha rhythm collapse predicts iso-electric suppressions during anesthesia. *Commun Biol* 2:327. <https://doi.org/10.1038/s42003-019-0575-3> PMID: 31508502; PMID: PMC6718680
35. Wong WEJ, Chan SP, Yong JK, Tham YYS, Lim JRG, Sim MA, Soh CR, Ti LK, Chew THS (2021) Assessment of acute kidney injury risk using a machine-learning guided generalized structural equation model: a cohort study. *BMC Nephrol* 22(1):63. <https://doi.org/10.1186/s12882-021-02238-9> PMID: 33618695; PMID: PMC7898752
36. Lee CK, Samad M, Hofer I, Cannesson M, Baldi P (2021) Development and validation of an interpretable neural network for prediction of postoperative in-hospital mortality. *NPJ Digit Med* 4(1):8. <https://doi.org/10.1038/s41746-020-00377-1> PMID: 33420341; PMID: PMC7794438
37. Jeong YS, Kim J, Kim D, Woo J, Kim MG, Choi HW, Kang AR, Park SY (2021) Prediction of postoperative complications for patients of end stage renal disease. *Sensors (Basel)* 21(2):544. <https://doi.org/10.3390/s21020544> PMID: 33466610; PMID: PMC7828737
38. Filiberto AC, Ozrazgat-Baslanti T, Loftus TJ, Peng YC, Datta S, Efron P, Upchurch GR Jr, Bihorac A, Cooper MA (2021) Optimizing predictive strategies for acute kidney injury after major vascular surgery. *Surgery* 170(1):298–303. <https://doi.org/10.1016/j.surg.2021.01.030> Epub 2021 Feb 27. PMID: 33648766; PMID: PMC8276529
39. Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, Sündermann SH, Stamm C, Hofmann T, Falk V, Eickhoff C (2018) Machine learning for real-time prediction of complications in critical care: a retrospective study. *Lancet Respir Med* 6(12):905–914. [https://doi.org/10.1016/S2213-2600\(18\)30300-X](https://doi.org/10.1016/S2213-2600(18)30300-X) Epub 2018 Sep 28. PMID: 30274956
40. Chiew CJ, Liu N, Wong TH, Sim YE, Abdullah HR (2020) Utilizing machine learning methods for preoperative prediction of postsurgical mortality and intensive care unit admission. *Ann Surg* 272(6):1133–1139. <https://doi.org/10.1097/SLA.0000000000003297> PMID: 30973386; PMID: PMC7668340
41. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, Motaei A, Madkour M, Pardalos PM, Lipori G, Hogan WR, Efron PA, Moore F, Moldawer LL, Wang DZ, Hobson CE, Rashidi P, Li X, Momcilovic P (2019) MySurgeryRisk: Development and validation of a machine-learning risk algorithm for major complications and death after surgery. *Ann Surg* 269(4):652–662. <https://doi.org/10.1097/SLA.0000000000002706> PMID: 29489489; PMID: PMC6110979
42. Yao RQ, Jin X, Wang GW, Yu Y, Wu GS, Zhu YB, Li L, Li YX, Zhao PY, Zhu SY, Xia ZF, Ren C, Yao YM (2020) A machine learning-based prediction of hospital mortality in patients with postoperative sepsis. *Front Med (Lausanne)*. 7:445. <https://doi.org/10.3389/fmed.2020.00445> PMID: 32903618; PMID: PMC7438711
43. Datta S, Loftus TJ, Ruppert MM, Giordano C, Upchurch GR Jr, Rashidi P, Ozrazgat-Baslanti T, Bihorac A (2020) Added value of intraoperative data for predicting postoperative complications: the MySurgeryRisk PostOp Extension. *J Surg Res* 254:350–363. <https://doi.org/10.1016/j.jss.2020.05.007> Epub 2020 Jun 9. PMID: 32531520; PMID: PMC7755426
44. Brennan M, Puri S, Ozrazgat-Baslanti T, Feng Z, Ruppert M, Hashemighouchani H, Momcilovic P, Li X, Wang DZ, Bihorac A (2019) Comparing clinical judgment with the MySurgeryRisk algorithm for preoperative risk assessment: a pilot usability study. *Surgery* 165(5):1035–1045. <https://doi.org/10.1016/j.surg.2019.01.002> Epub 2019 Feb 18. PMID: 30792011; PMID: PMC6502657
45. Houthoofd R, Ruysinck J, van der Hertten J, Stijven S, Couckuyt I, Gadeyne B, Ongenaet F, Colpaert K, Decruyenaere J, Dhaene T, De Turck F (2015) Predictive modelling of survival and length of stay in critically ill patients using sequential organ failure scores. *Artif Intell Med* 63(3):191–207. <https://doi.org/10.1016/j.artmed.2014.12.009> Epub 2014 Dec 30. PMID: 25579436
46. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, Liston DE, Low DK, Newman SF, Kim J, Lee SI (2018) Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng* 2(10):749–760. <https://doi.org/10.1038/s41551-018-0304-0> Epub 2018 Oct 10. PMID: 31001455; PMID: PMC6467492
47. Kang AR, Lee J, Jung W, Lee M, Park SY, Woo J, Kim SH (2020) Development of a prediction model for hypotension after induction of anesthesia using machine learning. *PLoS One* 15(4):e0231172. <https://doi.org/10.1371/journal.pone.0231172> PMID: 32298292; PMID: PMC7162491
48. Tan HS, Liu N, Sultana R, Han NR, Tan CW, Zhang J, Sia ATH, Sng BL (2021) Prediction of breakthrough pain during labour neuraxial analgesia: comparison of machine learning and multivariable regression approaches. *Int J Obstet Anesth* 45:99–110. <https://doi.org/10.1016/j.ijoa.2020.08.010> Epub 2020 Aug 25. PMID: 33121883
49. Shillan D, Sterne JAC, Champneys A, Gibbison B (2019) Use of machine learning to analyse routinely collected intensive care unit data: a systematic review. *Crit Care* 23(1):284. <https://doi.org/10.1186/s13054-019-2564-9> PMID: 31439010; PMID: PMC6704673
50. Van de Sande D, van Genderen ME, Huiskens J, Gommers D, van Bommel J (2021) Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit. *Intensive Care Med* 47(7): 750–760. <https://doi.org/10.1007/s00134-021-06446-7> Epub 2021 Jun 5. PMID: 34089064; PMID: PMC8178026
51. Connor CW (2019) Artificial intelligence and machine learning in anesthesiology. *Anesthesiology* 131(6):1346–1359. <https://doi.org/10.1097/A.LN.0000000000002694> PMID: 30973516; PMID: PMC6778496
52. Gambus P, Shafer SL. (2018). Artificial Intelligence for Everyone. *Anesthesiology*;128(3):431-433. doi: <https://doi.org/10.1097/ALN.0000000000001984>. PMID: 29166324.
53. Liu Q, Ma L, Fan SZ, Abbd MF, Lu CW, Lin TY, Jen KK, Wu SJ, Shieh JS. (2018). Design and evaluation of a real time physiological signals acquisition system implemented in multi-operating rooms for anesthesia. *J Med Syst*; 42(8):148. doi: <https://doi.org/10.1007/s10916-018-0999-1>. PMID: 29961144.
54. Chakravarthy M (2017) Modifying risks to improve outcome in cardiac surgery: an anesthesiologist's perspective. *Ann Card Anaesth* 20(2):226–233. https://doi.org/10.4103/aca.ACA_20_17 PMID: 28393785; PMID: PMC5408530
55. Wang Z, Majewicz FA (2018) Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. *Int J Comput Assist Radiol Surg* 13(12):1959–1970. <https://doi.org/10.1007/s11548-018-1860-1> Epub 2018 Sep 25. PMID: 30255463
56. Rahmani, A.M.; Yousefpoor, E.; Yousefpoor, M.S.; Mehmood, Z.; Haider, A.; Hosseinzadeh, M.; Ali Naqvi, R. (2021). Machine learning (ML) in medicine: review, applications, and challenges. *Mathematics*; 9, 2970. <https://doi.org/10.3390/math9222970>
57. Bellini V, Petroni A, Palumbo G, Bignami E (2019) Data quality and blockchain technology. *Anaesth Crit Care Pain Med* 38(5):521–522. <https://doi.org/10.1016/j.jaccpm.2018.12.015> Epub 2019 Jan 8. PMID: 30633992
58. Collins GS, Moons KGM. (2019). Reporting of artificial intelligence prediction models. *Lancet*;393(10181):1577-1579. doi: [https://doi.org/10.1016/S0140-6736\(19\)30037-6](https://doi.org/10.1016/S0140-6736(19)30037-6). PMID: 31007185.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.