

Random Forest-Based Prediction of Acute Respiratory Distress Syndrome in Patients Undergoing Cardiac Surgery

Wei Wang, MD,¹ Lina Li, MD,¹ Hongjun Gu, MD,² Yanqing Chen, MD,³ Yumei Zhen, MD,⁴ Zhaorui Dong, MD⁵

¹Department of Respiratory and Critical Care Medicine, the First Medical Centre of Chinese PLA General Hospital, Haidian District, Beijing China; ²Department of Respiratory and Critical Care Medicine, the Eighth Medical Centre of Chinese PLA General Hospital, Haidian District, Beijing China; ³Department of Haematology, the Second Medical Centre of Chinese PLA General Hospital, Haidian District, Beijing China; ⁴Department of Rehabilitation, Xinkang hospital, Daxing District, Beijing China; ⁵Department of Cardiovascular Surgery, the First Medical Centre of Chinese PLA General Hospital, Haidian District, Beijing China

ABSTRACT

Background: To develop a machine learning-based model for predicting the risk of acute respiratory distress syndrome (ARDS) after cardiac surgery.

Methods: Data were collected from 1011 patients, who underwent cardiac surgery between February 2018 and September 2019. We developed a predictive model on ARDS by using the random forest algorithm of machine learning. The discrimination of the model was then shown by the area under the curve (AUC) of the receiver operating characteristic curve. Internal validation was performed by using a 5-fold cross-validation technique, so as to evaluate and optimize the predictive model. Model visualization was performed to reveal the most influential features during the model output.

Results: Of the 1011 patients included in the study, 53 (5.24%) suffered ARDS episodes during the first postoperative week. This random forest distinguished ARDS patients from non-ARDS patients with an AUC of 0.932 (95% CI=0.896-0.968) in the training set and 0.864 (95% CI=0.718-0.997) in the final test set. The top 10 variables in the random forest were cardiopulmonary bypass time, transfusion red blood cell, age, EuroSCORE II score, albumin, hemoglobin, operation time, serum creatinine, diabetes, and type of surgery.

Conclusion: Our findings suggest that machine learning algorithm is highly effective in predicting ARDS in patients undergoing cardiac surgery. The successful application of the generated random forest may guide clinical decision-making and aid in improving the long-term prognosis of patients.

INTRODUCTION

Acute respiratory distress syndrome (ARDS) significantly compromises the prognosis of patients subjected to cardiac surgery, with a mortality rate of up to 40% [Zhang 2021]. Despite the availability of multiple interventions, such as pulmonary protective mechanical ventilation, fluid management, glucocorticoid administration, and other measures to preserve organ function, effective treatments for ARDS remain relatively limited [Bricher Choque 2021; Liu 2021]. As such, most studies on ARDS are currently aimed at achieving early detection and taking effective measures to avert its occurrence, thereby improving the outcome of patients. With identification of risk factors and establishment of predictive scores, one can identify patients at high risk of ARDS early and take preventive measures before the onset of ARDS, thus reducing the morbidity and mortality of ARDS. However, there are few available studies on establishing ARDS prediction scores. Currently, the outcome of cardiac surgery markedly varies based on the type of procedure, intraoperative blood transfusion, and postoperative factors, which all affect the development of ARDS [Su 2019]. To our knowledge, there are no machine learning models available that are exclusively designed to predict the occurrence of ARDS after cardiac surgery. The aim of our study is to identify high-risk ARDS patients early by machine learning to assist clinicians in making decisions and taking early precautions.

METHODS

Study design and population: Analyzed data from patients, who underwent adult cardiac surgery between February 2018 and September 2019, were extracted from one medical center, the General Hospital of the Chinese PLA General Hospital. The study was approved by the local Institutional Review Board, which waived informed consent, due to the observational nature of the study. The study was reported in accordance with the recommendations of the Reporting of Observational Studies in Epidemiology (STROBE) criteria for observational studies.

Received September 5, 2022; accepted October 3, 2022.

Correspondence: Dr. Zhaorui Dong, Department of Cardiovascular Surgery, the First Medical Centre of Chinese PLA General Hospital, No. 28, Fuxing Road, Haidian District, 100853, Beijing, China (e-mail: DongZR301hospital@163.com).

Figure 1 shows the flow chart describing the study protocol. (Figure 1) In all, 1145 consecutive patients were selected from the electronic medical record system between February 2018 and September 2019. Exclusion criteria were 1) age <18 years, and 2) over 10% loss of perioperative data. The final inclusion of 1011 patients comprised the entire dataset. In the present study, all patients enrolled in the dataset were grouped into two subsets: 70% were categorized into a training subset for the training of the RF model, and the remaining 30% were employed as a test subset to validate the RF model performance.

A total of 1145 patients, who underwent cardiac surgery, were preliminarily recruited. According to the exclusion criteria, 134 patients were excluded, leaving a total of 1011 patients comprised the final dataset. The finished dataset was split into a training set and a test set in a ratio of 7 to 3. The random forest was then developed on the training set and further validated in the test set.

Data collection: Table 1 summarizes the preoperative and perioperative variables collected from the electronic health record. (Table 1) 1) Preoperative variables included age, gender, body mass index (BMI), EuroSCORE II Score, NYHA functional classification, medical history, and preoperative condition. Laboratory findings concerned hemoglobin, white blood cell, alanine aminotransferase, aspartate aminotransferase, albumin, and serum creatinine also were collected. 2) Intraoperative variables such as type of surgery, duration of surgery, cardiopulmonary bypass (CPB) time, perioperative blood loss, and transfused red blood cells also were included in the analysis.

Primary outcome: The definition of ARDS follows the current Berlin criteria [Ball 2022]. It is characterized by: (1) A known clinical insult or newly present or worsening symptoms within 1 week; (2) Bilateral opacities on chest imaging that cannot be fully interpreted as exudates, lobar/lung collapse, or nodules; and (3) An edematous origin that cannot be fully interpreted as heart failure or fluid overload. Hypoxia was defined as $200 \text{ mmHg} < \text{partial pressure of alveolar oxygen (PAO}_2\text{)}/\text{partial pressure of inspired oxygen (FiO}_2\text{)} \leq 300 \text{ mmHg}$, with positive end-expiratory pressure (PEEP) or continuous positive airway pressure (CPAP) $\geq 5 \text{ cmH}_2\text{O}$.

Random forest model: Random forest, as a widely used machine learning model, is a non-parametric and supervised ensemble machine learning tool that originally was proposed by Breiman as an extension to solve classification and regression problems [Yuan 2019]. The random forest is based on methods that train a forest of binary decision trees, where Fisher's discriminant is used as a linear classifier for each branch of the tree. In an ensemble decision tree, the algorithm employs a binary arithmetic approach to split the observations into two homologous groups, called branches, and repeats this splitting process until the "tree" is completely grown ["node purity" is reached].

During the random forest modelling, the entire dataset was classified into two subsets: 70% for model training and 30% for model testing. In the training set, the hyper-parameters of the model were obtained through a 5-fold cross-validation process. On a more specific note, 70% of the training

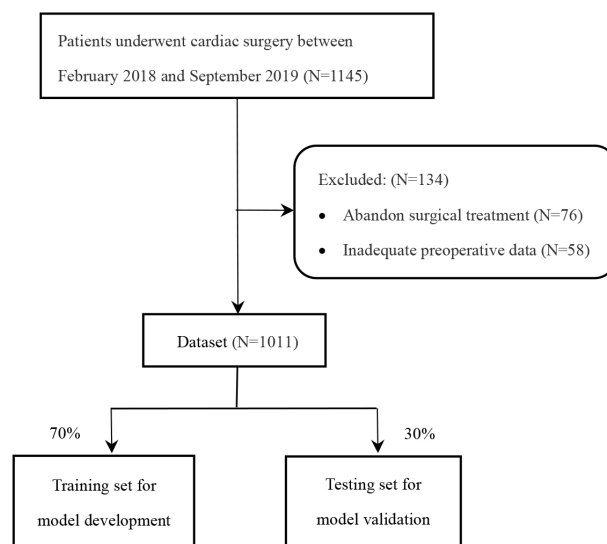


Figure 1. Flow chart presenting patient selection and analysis

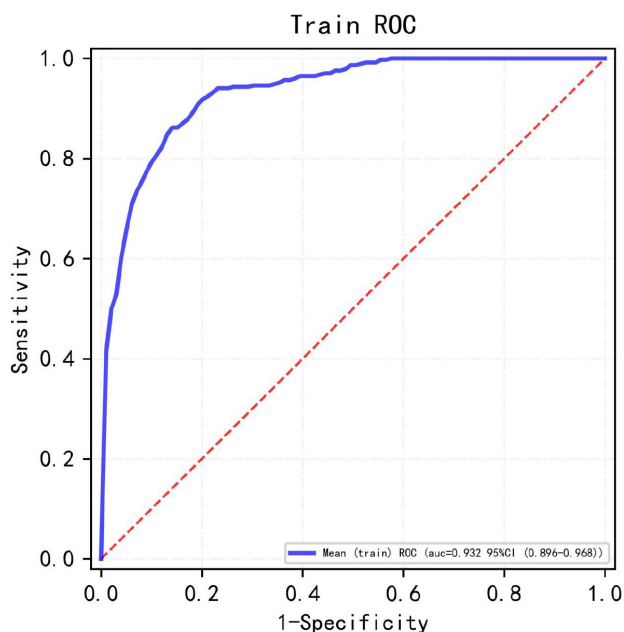


Figure 2. AUC of the ROC of random forest model in the train set. AUC, area under curve; ROC, receiver operating characteristic

set randomly was divided into 80% and 20% sub-group for model training and testing, respectively. The 5-fold cross-validation is to perform this cross-testing 5 times, thus obtaining 5 individual scores. And the ultimate tuned hyper-parameters of the model were the average of these 5 individual scores. Finally, the random forest model developed from the training set were fed into the remaining 30% test set to validate the model performance.

In this study, we employed the most prevalent metric to evaluate the performance of random forest model, using the

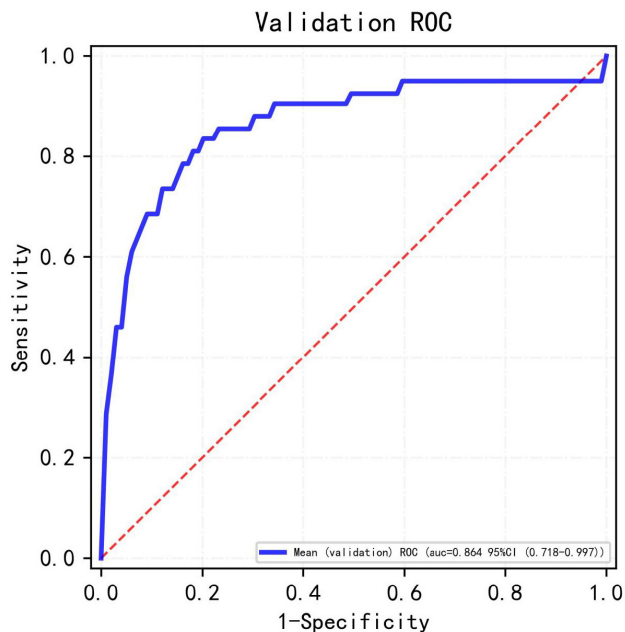


Figure 3. AUC of the ROC of random forest model in the test set. AUC, area under curve; ROC, receiver operating characteristic

receiver operating characteristic curve to determine the area under the curve (AUC): the greater the AUC, the superior the predictive model.

Statistical analysis: For data analyses, packages were implemented using Python software (version 3.6) and Scikit-learn (<https://scikit-learn.org/>). Descriptive statistics were expressed as medians (interquartile range) or numbers (percentages), and statistical analyses were conducted using the Mann-Whitney U test or Pearson chi-square test, as appropriate. $P < 0.05$ was considered as the cut-off value for statistical significance.

RESULTS

Patient characteristics: Data pertaining to 1011 consecutive patients who underwent cardiac surgery between February 2018 and September 2019 constituted the entire dataset. The population had a median age (interquartile range) of 58 years (49-65) at the time of surgery, a median BMI of 24.95 kg/m² (22.40-27.00), and a median EuroSCORE II score of 1.52 (0.84-2.88). Among them, 59.84% (605/1011) were male, 34.72% (351/1011) were smokers, 3.86% (39/1011) had infectious endocarditis, and 3.17% (32/1011) had myocardial infarction within 90 days. Notably, 58.06% (587/1011), 28.68% (290/1011), 6.73% (68/1011), and 6.53% (66/1011) of patients underwent valve surgery, coronary artery bypass grafting (CABG) surgery, CABG plus valve surgery, and aortic surgery, respectively (Table 1).

Postoperative ARDS morbidity: According to the definition of Berlin criteria, out of 1011 patients in this study, a total of 53 (5.24%) suffered ARDS episodes within 7 days

after surgery. Table 1 outlines the relevant data comparing patients who developed ARDS and those who did not. In summary, there was a statistically significant ($P < 0.05$) difference between the ARDS and non-ARDS groups in both preoperative and intraoperative variables (Table 1).

Model performance: To have in-depth insight into the performance of the random forest model, we performed receiver operating characteristic (ROC) curve analysis, as it considers both sensitivity and specificity, while the area under the curve is considered a valid measure of accuracy. In the ROC curve, when calculated by random prediction, the value of AUC is 0.50, while an AUC value of 1 represents a 100% discrimination. In general, the higher the value of AUC, the better the performance of the model, and an AUC > 0.8 indicates a model with high discrimination ability. In the present study, the random forest showed high discriminative power, in terms of the model's effectiveness in predicting ARDS, with an AUC value of 0.932 (95% CI=0.896-0.968) in the training set. (Figure 2) In the final test set, the model consistently displayed a strong discriminative power, with its AUC value of 0.864 (95% CI=0.718-0.997). (Figure 3)

Variable importance ranking: In this study, an interpretable machine learning algorithm is attempted to visualize the importance ranking of variables in the random forest model. As shown in Figure 1, the top 10 ranked risk factors include cardiopulmonary bypass time, transfusion red blood cell, age, EuroSCORE II score, albumin, hemoglobin, operation time, serum creatinine, diabetes, and type of surgery. Notably, among the top 10 variables, there are four intraoperative variables, namely, cardiopulmonary bypass time, transfusion red blood cell, operation time, and type of surgery. (Figure 4)

DISCUSSION

In this study, we present an approach allowing early prediction of ARDS onset after cardiac surgery using a supervised machine learning model - the random forest. In our validation, the random forest model performed well in predicting postoperative ARDS, according to the Berlin definition. The high AUC of the model demonstrates its utility in early identification of patients at high risk of developing ARDS. We developed this model using preoperative and intraoperative clinical factors extracted from patient's electrical health records. Specifically, we visualized the top 10 factors of the model in predicting ARDS after cardiac surgery, including cardiopulmonary bypass time, transfusion red blood cell, age, EuroSCORE II score, albumin, hemoglobin, operation time, serum creatinine, diabetes, and type of surgery.

As a non-cardiac pulmonary edema, ARDS is an alveolar injury caused by inflammation, which clinically presented as an acute outbreak of bilateral infiltrates (evident on chest radiograph) together with arterial hypoxemia [Liang 2012; Zhang 2022]. Patients undergoing cardiac surgery also are at high risk for ARDS. In fact, ARDS has been reported in 0.4%-8.1% of patients after cardiac surgery, which is associated with increased in-hospital mortality and hospital costs [Teixeira 2019; Sanfilippo 2022]. Of the 1011 patients

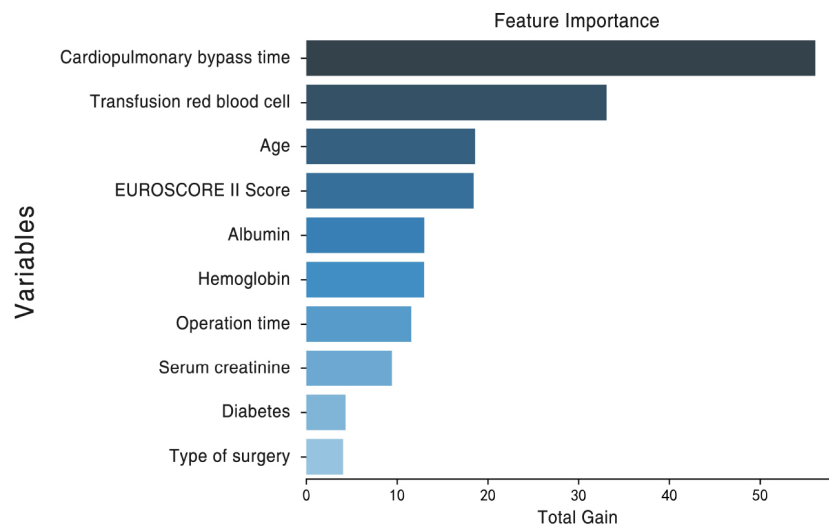


Figure 4. Feature importance ranking of the random forest. The importance ranking plot depicts the top 10 features in the development of the final predictive model.

included in the present study, a total of 53 patients suffered ARDS episodes within 7 days after surgery. The incidence of postoperative ARDS was 5.24%, which is consistent with that reported in previous literature.

Despite the fact that mechanical ventilation strategies are proven to affect mortality in ARDS [Yener 2020], the failure to foresee which patients may develop ARDS poses a major challenge for early intervention and prevention. To allow early identification of patients at risk of developing ARDS after cardiac surgery, several risk scores based on multivariable logistic regression have been developed [Huang 2021; Liu 2021]. For example, Huang et al. derived an ARDS prediction score from a retrospective derivation cohort and then further validated it in a prospective cohort. In this study, discrimination was assessed using the AUC metric, and its value in the validation cohort was 0.78 (95% CI, 0.71-0.85) [Huang 2021]. In addition, in another retrospective study, a nomogram to predict ARDS after cardiac surgery was developed by using multivariable logistic regression. And the AUC of the nomogram to distinguish ARDS patients from non-ARDS patients was 0.785 (95% CI: 0.740, 0.830) [Liu 2021]. In summary, the previously developed, logistic regression-based classifiers showed only moderate accuracy (AUC of 0.75-0.80) during the validation process. Nevertheless, in the present study, we developed a random forest-based machine learning model, with an AUC of 0.864 (95% CI=0.718-0.997). Several factors may have contributed to the better performance of the model presented in this study than models developed in previous studies. First, the model proposed in this study may benefit from the higher prediction accuracy that comes from using random forests. Unlike the widely established logistic regression, the random forest, as an ensemble of weak predictive models, has better performance when dealing with high-dimensional data [Hajipour 2020; Ramos 2021]. It is capable of capturing potential nonlinear

relationships between variables and outcomes during modeling [Nath 2022]. In addition, compared with traditional logistic regression, random forest models have the capability of optimizing hyperparameters through cross-validation and grid search functions. This gives random forest a clear advantage to construct optimal models by multiple internal validations [Scott 2020]. Furthermore, it is noteworthy that not only preoperative variables but also surgery-related variables were included in the modeling process. This enables better simulation of the real physiological conditions during cardiac surgery [Dong 2021]. Consequently, if there is heterogeneity among patients' conditions, random forest could offer better detection of such differences.

Given the performance of the obtained random forest model, results from our study have important clinical application in early detection of ARDS. Our model showed favorable predictive capacity for ARDS, with an AUC score of 0.864 (95% CI=0.718-0.997). This property enables clinicians to instantly discern patients at high risk for ARDS, particularly those with severe forms of ARDS who may require mechanical ventilation and other advanced therapies. As delayed intubation is known to increase mortality in ARDS, early identification of the severity of ARDS and eligibility for invasive mechanical ventilation is critical for later survival [Kangelaris 2016]. Notably, such early risk stratification offers a "second opinion" for decision-making, such as the timing of intubation in critically ill patients. The random forest model in this study could detect the development of ARDS by analyzing patient preoperative and intraoperative variables, even before the actual onset of the disease, thus warning clinicians of patients who are at risk of developing ARDS potentially and urging them to assess the necessity of intubation earlier. Furthermore, early identifying high-risk patients allows for the timely intervention of evidence-based strategies to avoid further deterioration. These treatment strategies include

low tidal volume and lung-protective ventilation tactics for patients already receiving mechanical ventilation [Wong 2020], fluid balance and early use of diuretics [Seitz 2020]. In the clinic, our prediction model may allow more time to preempt these proven strategies to mitigate lung injury in patients with progressively worsening hypoxia, thus preclude the progression of ARDS and improve later prognosis.

The feature importance ranking in our model offers clues to recognize the most important clinical features for the onset of ARDS. Not surprisingly, features that are directly associated with the surgical procedure of the patient, including cardiopulmonary bypass time, transfusion red blood cell, operation time and type of surgery, were measured as the top 10 variables. Specifically, among the top 10 variables, cardiopulmonary bypass time was measured as the most important feature in predicting ARDS. This result is consistent with the findings of previous logistic regression, in which cardiopulmonary bypass time also was identified as a significant predictor in detecting ARDS [Huang 2021; Liu 2021]. Epidemiological studies also have shown that cardiac surgery with cardiopulmonary bypass is a well-known risk factor for ARDS, and that cardiopulmonary bypass surgery often is involved with lung injury, and sometimes with ARDS [Stephens 2013]. The exposure of blood to aphysiological surfaces, the concomitant ischemia-reperfusion injury, and the transfer of endotoxins from the gut into the bloodstream can potentially activate multiple inflammatory pathways, leading to a systemic inflammatory response [Cheng 2020]. In the present machine learning model, the results also show that cardiopulmonary bypass time possesses indelible contributions to alerting ARDS, with highly discriminatory power in discerning disease sub-phenotype.

This study, necessarily, is subject to several limitations. The retrospective nature of the study subjects it to selection bias and makes it impossible to assert causality. Further, the variables employed to construct the models were collected retrospectively. Thus, the performance of the model in a directly extracted real-time data environment also is currently unknown, which may pose a significant challenge to the implementation of the model at the bedside. In addition, we have not yet demonstrated the performance of more recent deep learning methods. Incorporating these algorithms may further improve the performance of our models. Furthermore, this was a single-centered study with a relatively small sample size of 1011 patients, which may have limited the generalizability of our findings. Lastly, another limitation of our study is the utilization of internal validation, which has less power than external validation with a prospective population. Therefore, additional prospective studies with larger sample sizes are warranted to further evaluate our model.

CONCLUSION

ARDS, a severe form of acute lung injury, is a devastating complication that can occur after cardiac surgery, which is associated with significant mortality and prolonged ventilation. In this study, we employed random forest to predict

the development of ARDS in patients undergoing cardiac surgery. The results show that random forest model possesses high predictive power in predicting ARDS after cardiac surgery. To our knowledge, this is the first demonstration of machine learning in predicting ARDS after cardiac surgery. In the context of personalized medicine, we believe that machine learning will enable us to assess the risk of ARDS accurately and then take early preventive measures for those patients at high risk.

REFERENCES

- Ball L, Silva PL, Giacobbe DR, Bassetti M, Zubieta-Calleja GR, Rocco P, et al. 2022. Understanding the pathophysiology of typical acute respiratory distress syndrome and severe COVID-19. *Expert Rev Respir Med.* 16(4):437-446.
- Bricher Choque PN, Vieira RP, Ulloa L, Grabulosa C, Irigoyen MC, De Angelis K, et al. 2021. The Cholinergic Drug Pyridostigmine Alleviates Inflammation During LPS-Induced Acute Respiratory Distress Syndrome. *Front Pharmacol.* 12:624895.
- Cheng ZB, Chen H. 2020. Higher incidence of acute respiratory distress syndrome in cardiac surgical patients with elevated serum procalcitonin concentration: a prospective cohort study. *Eur J Med Res.* 25(1):11.
- Dong JF, Xue Q, Chen T, Zhao YY, Fu H, Guo WY, et al. 2021. Machine learning approach to predict acute kidney injury after liver surgery. *World J Clin Cases.* 9(36):11255-11264.
- Hajipour F, Jozani MJ, Moussavi Z. 2020. A comparison of regularized logistic regression and random forest machine learning models for daytime diagnosis of obstructive sleep apnea. *Med Biol Eng Comput.* 58(10):2517-2529.
- Huang L, Song M, Liu Y, Zhang W, Pei Z, Liu N, et al. 2021. Acute Respiratory Distress Syndrome Prediction Score: Derivation and Validation. *Am J Crit Care.* 30(1):64-71.
- Kangelaris KN, Ware LB, Wang CY, Janz DR, Zhuo H, Matthay MA, et al. 2016. Timing of Intubation and Clinical Outcomes in Adults With Acute Respiratory Distress Syndrome. *Crit Care Med.* 44(1):120-129.
- Liang Y, Yeligar SM, Brown LA. 2012. Chronic-alcohol-abuse-induced oxidative stress in the development of acute respiratory distress syndrome. *Scientific World Journal.* 2012:740308.
- Liu J, Sala MA, Kim J. 2021. Dampening the Fire: A Negative Feedback Loop in Acute Respiratory Distress Syndrome. *Am J Respir Cell Mol Biol.* 64(2):158-160.
- Liu Y, Song M, Huang L, Zhu G. 2021. A Nomogram to Predict Acute Respiratory Distress Syndrome After Cardiac Surgery. *Heart Surg Forum.* 24(3):E445-E450.
- Nath B, Chowdhury R, Ni-Meister W, Mahanta C. 2022. Predicting the Distribution of Arsenic in Groundwater by a Geospatial Machine Learning Technique in the Two Most Affected Districts of Assam, India: The Public Health Implications. *Geohealth.* 6(3):e2021GH000585.
- Ramos LA, Blankers M, van Wingen G, de Bruijn T, Pauws SC, Goudriaan AE. 2021. Predicting Success of a Digital Self-Help Intervention for Alcohol and Substance Use With Machine Learning. *Front Psychol.* 12:734633.
- Sanfilippo F, Palumbo GJ, Bignami E, Pavesi M, Ranucci M, Scolletta S, et al. 2022. Acute Respiratory Distress Syndrome in the Perioperative Period of Cardiac Surgery: Predictors, Diagnosis, Prognosis,

- Management Options, and Future Directions. *J Cardiothorac Vasc Anesth.* 36(4):1169-1179.
- Scott CA, Duryea JD, MacKay H, Baker MS, Laritsky E, Gunasekara CJ, et al. 2020. Identification of cell type-specific methylation signals in bulk whole genome bisulfite sequencing data. *Genome Biol.* 21(1):156.
- Seitz KP, Caldwell ES, Hough CL. 2020. Fluid management in ARDS: an evaluation of current practice and the association between early diuretic use and hospital mortality. *J Intensive Care.* 8:78.
- Stephens RS, Shah AS, Whitman GJ. 2013. Lung injury and acute respiratory distress syndrome after cardiac surgery. *Ann Thorac Surg.* 95(3):1122-1129.
- Su IL, Wu VC, Chou AH, Yang CH, Chu PH, Liu KS, et al. 2019. Risk factor analysis of postoperative acute respiratory distress syndrome after type A aortic dissection repair surgery. *Medicine (Baltimore).* 98(29):e16303.
- Teixeira C, Rosa RG, Maccari JG, Savi A, Rotta FT. 2019. Association between electromyographical findings and intensive care unit mortality among mechanically ventilated acute respiratory distress syndrome patients under profound sedation. *Rev Bras Ter Intensiva.* 31(4):497-503.
- Wong J, Lee SW, Tan HL, Ma YJ, Sultana R, Mok YH, et al. 2020. Lung-Protective Mechanical Ventilation Strategies in Pediatric Acute Respiratory Distress Syndrome. *Pediatr Crit Care Med.* 21(8):720-728.
- Yener N, Üdügücü M. 2020. Airway Pressure Release Ventilation as a Rescue Therapy in Pediatric Acute Respiratory Distress Syndrome. *Indian J Pediatr.* 87(11):905-909.
- Yuan H, Fan XS, Jin Y, He JX, Gui Y, Song LY, et al. 2019. Development of heart failure risk prediction models based on a multi-marker approach using random forest algorithms. *Chin Med J (Engl).* 132(7):819-826.
- Zhang C, Huang Q, He F. 2022. Correlation of small nucleolar RNA host gene 16 with acute respiratory distress syndrome occurrence and prognosis in sepsis patients. *J Clin Lab Anal.* 36(7):e24516.
- Zhang R, Chen H, Gao Z, Liang M, Qiu H, Yang Y, et al. 2021. The Effect of Loop Diuretics on 28-Day Mortality in Patients With Acute Respiratory Distress Syndrome. *Front Med (Lausanne).* 8:740675.