

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

Wei Wang¹, Lina Li¹, Hongjun Gu¹, Yanqing Chen¹, Yumei Zhen², and Zhaorui Dong¹

¹Chinese PLA General Hospital

²Department of Rehabilitation Xinkang hospital Daxing District Beijing China

July 25, 2022

Abstract

Objective: 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.

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 and 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

Correspondence: Zhaorui Dong, MD, Department of Cardiovascular Surgery, the First Medical Centre of Chinese PLA General Hospital, No. 28, Fuxing Road, Haidian District, 100853, Beijing, China. Email:

DongZR301hospital@163.com

Abstract

Objective: 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.

Keywords: cardiac surgery, machine learning, ARDS, predictive model

Abbreviations

ARDS acute respiratory distress syndrome

AUC area under the curve

BMI body mass index

CABG coronary artery bypass grafting

CPAP continuous positive airway pressure

CPB cardiopulmonary bypass

FiO₂ partial pressure of inspired oxygen

PAO₂ partial pressure of alveolar oxygen

PEEP positive end-expiratory pressure

INTRODUCTION

Acute respiratory distress syndrome (ARDS) significantly compromises the prognosis of patients subjected to cardiac surgery, with a mortality rate of up to 40%.¹ 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.^{2,3} 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.⁴ 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.

Figure 1 shows the flow chart describing the study protocol. 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 2 sub-sets: 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.

Data collection

Table 1 summarizes the preoperative and perioperative variables collected from the electronic health record. 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 were also collected. 2) Intraoperative variables such as type of surgery, duration of surgery, cardiopulmonary bypass (CPB) time, perioperative blood loss, and transfused red blood cells were also included in the analysis.

Primary Outcome

The definition of ARDS follows the current Berlin criteria.⁵ 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. (4) 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, which originally proposed by Breiman as an extension to solve classification and regression problems.⁶ 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 employ 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 set was randomly 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 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 between patients who developed ARDS and those who did not. In summary, there were 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 an 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 4 intra-operative 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 post-operative 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.^{7,8} Patients undergoing cardiac surgery are also 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.^{9,10} Of the 1011 patients 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,¹¹ 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.^{12,13} For example, Huang and colleagues 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).¹² 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).¹³ 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.^{14,15} It is capable of capturing potential nonlinear relationships between variables and outcomes during modeling.¹⁶ In addition, compared to 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.¹⁷ 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.¹⁸ 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.¹⁹ 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,²⁰ fluid balance and early use of diuretics.²¹ 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 was also identified as a significant predictor in detecting ARDS.^{12,13} Epidemiological studies have also shown that cardiac surgery with cardiopulmonary bypass is a well-known risk factor for ARDS, and that cardiopulmonary bypass surgery is often involved with lung injury, and sometimes with ARDS.²² The exposure of blood to aphysiological surfaces, the concomitant ischaemia-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.²³ 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 is also 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 singly-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.

CONFLICT OF INTERESTS

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ETHICS STATEMENT

The study was approved by the local Institutional Review Board, which waived informed consent due to the observational nature of the study.

REFERENCES

1. Zhang R, Chen H, Gao Z, et al. The Effect of Loop Diuretics on 28-Day Mortality in Patients With Acute Respiratory Distress Syndrome. *Front Med (Lausanne)* . 2021;8:740675. <https://doi.org/10.3389/fmed.2021.740675>
2. Liu J, Sala MA, Kim J. Dampening the Fire: A Negative Feedback Loop in Acute Respiratory Distress Syndrome. *Am J Respir Cell Mol Biol* . 2021;64(2):158-160. <https://doi.org/10.1165/rcmb.2020-0487ED>

3. Bricher Choque PN, Vieira RP, Ulloa L, et al. The Cholinergic Drug Pyridostigmine Alleviates Inflammation During LPS-Induced Acute Respiratory Distress Syndrome. *Front Pharmacol* . 2021;12:624895. <https://doi.org/10.3389/fphar.2021.624895>
4. Su IL, Wu VC, Chou AH, et al. Risk factor analysis of postoperative acute respiratory distress syndrome after type A aortic dissection repair surgery. *Medicine (Baltimore)* . 2019;98(29):e16303. <https://doi.org/10.1097/MD.00000000000016303>
5. Ball L, Silva PL, Giacobbe DR, et al. Understanding the pathophysiology of typical acute respiratory distress syndrome and severe COVID-19. *Expert Rev Respir Med* . 2022;16(4):437-446. <https://doi.org/10.1080/17476348.2022.2057300>
6. Yuan H, Fan XS, Jin Y, et al. Development of heart failure risk prediction models based on a multi-marker approach using random forest algorithms. *Chin Med J (Engl)* . 2019;132(7):819-826. <https://doi.org/10.1097/CM9.0000000000000149>
7. Zhang C, Huang Q, He F. Correlation of small nucleolar RNA host gene 16 with acute respiratory distress syndrome occurrence and prognosis in sepsis patients. *J Clin Lab Anal* . 2022;36(7):e24516. <https://doi.org/10.1002/jcla.24516>
8. Liang Y, Yeligar SM, Brown LA. Chronic-alcohol-abuse-induced oxidative stress in the development of acute respiratory distress syndrome. *ScientificWorldJournal* . 2012;2012:740308. <https://doi.org/10.1100/2012/740308>
9. Sanfilippo F, Palumbo GJ, Bignami E, et al. Acute Respiratory Distress Syndrome in the Perioperative Period of Cardiac Surgery: Predictors, Diagnosis, Prognosis, Management Options, and Future Directions. *J Cardiothorac Vasc Anesth* . 2022;36(4):1169-1179. <https://doi.org/10.1053/j.jvca.2021.04.024>
10. Teixeira C, Rosa RG, Maccari JG, Savi A, Rotta FT. Association between electromyographical findings and intensive care unit mortality among mechanically ventilated acute respiratory distress syndrome patients under profound sedation. *Rev Bras Ter Intensiva* . 2019;31(4):497-503. <https://doi.org/10.5935/0103-507X.20190087>
11. Yener N, Üdürücü M. Airway Pressure Release Ventilation as a Rescue Therapy in Pediatric Acute Respiratory Distress Syndrome. *Indian J Pediatr* . 2020;87(11):905-909. <https://doi.org/10.1007/s12098-020-03235-w>
12. Huang L, Song M, Liu Y, et al. Acute Respiratory Distress Syndrome Prediction Score: Derivation and Validation. *Am J Crit Care* . 2021;30(1):64-71. <https://doi.org/10.4037/ajcc2021753>
13. Liu Y, Song M, Huang L, Zhu G. A Nomogram to Predict Acute Respiratory Distress Syndrome After Cardiac Surgery. *Heart Surg Forum* . 2021;24(3):E445-E450. <https://doi.org/10.1532/hsf.3809>
14. Ramos LA, Blankers M, van Wingen G, de Bruijn T, Pauws SC, Goudriaan AE. Predicting Success of a Digital Self-Help Intervention for Alcohol and Substance Use With Machine Learning. *Front Psychol* . 2021;12:734633. <https://doi.org/10.3389/fpsyg.2021.734633>
15. Hajipour F, Jozani MJ, Moussavi Z. A comparison of regularized logistic regression and random forest machine learning models for daytime diagnosis of obstructive sleep apnea. *Med Biol Eng Comput* . 2020;58(10):2517-2529. <https://doi.org/10.1007/s11517-020-02206-9>
16. Nath B, Chowdhury R, Ni-Meister W, Mahanta C. 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* . 2022;6(3):e2021GH000585. <https://doi.org/10.1029/2021GH000585>
17. Scott CA, Duryea JD, MacKay H, et al. Identification of cell type-specific methylation signals in bulk whole genome bisulfite sequencing data. *Genome Biol* . 2020;21(1):156. <https://doi.org/10.1186/s13059-020-02065-5>

18. Dong JF, Xue Q, Chen T, et al. Machine learning approach to predict acute kidney injury after liver surgery. *World J Clin Cases* . 2021;9(36):11255-11264. <https://doi.org/10.12998/wjcc.v9.i36.11255>
19. Kangelaris KN, Ware LB, Wang CY, et al. Timing of Intubation and Clinical Outcomes in Adults With Acute Respiratory Distress Syndrome. *Crit Care Med* . 2016;44(1):120-129. <https://doi.org/10.1097/CCM.0000000000001359>
20. Wong J, Lee SW, Tan HL, et al. Lung-Protective Mechanical Ventilation Strategies in Pediatric Acute Respiratory Distress Syndrome. *Pediatr Crit Care Med* . 2020;21(8):720-728. <https://doi.org/10.1097/PCC.0000000000002324>
21. Seitz KP, Caldwell ES, Hough CL. Fluid management in ARDS: an evaluation of current practice and the association between early diuretic use and hospital mortality. *J Intensive Care* . 2020;8:78. <https://doi.org/10.1186/s40560-020-00496-7>
22. Stephens RS, Shah AS, Whitman GJ. Lung injury and acute respiratory distress syndrome after cardiac surgery. *Ann Thorac Surg* . 2013;95(3):1122-1129. <https://doi.org/10.1016/j.athoracsur.2012.10.024>
23. Cheng ZB, Chen H. Higher incidence of acute respiratory distress syndrome in cardiac surgical patients with elevated serum procalcitonin concentration: a prospective cohort study. *Eur J Med Res* . 2020;25(1):11. <https://doi.org/10.1186/s40001-020-00409-2>

Hosted file

Figure 1.docx available at <https://authorea.com/users/497155/articles/578312-random-forest-based-prediction-of-acute-respiratory-distress-syndrome-in-patients-undergoing-cardiac-surgery>

Hosted file

Figure 2.docx available at <https://authorea.com/users/497155/articles/578312-random-forest-based-prediction-of-acute-respiratory-distress-syndrome-in-patients-undergoing-cardiac-surgery>

Hosted file

Figure 3.docx available at <https://authorea.com/users/497155/articles/578312-random-forest-based-prediction-of-acute-respiratory-distress-syndrome-in-patients-undergoing-cardiac-surgery>

Hosted file

Figure 4.docx available at <https://authorea.com/users/497155/articles/578312-random-forest-based-prediction-of-acute-respiratory-distress-syndrome-in-patients-undergoing-cardiac-surgery>

Hosted file

Table. 1.docx available at <https://authorea.com/users/497155/articles/578312-random-forest-based-prediction-of-acute-respiratory-distress-syndrome-in-patients-undergoing-cardiac-surgery>