Posted on 24 Aug 2024 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.172448181.17701053/v1 | This is a preprint and has not been peer-reviewed. Data may be preliminary.

Early Prediction of CPAP Failure in Very Preterm Neonates Utilizing Machine Learning

Quinn Gates[1], Shruthi Janardhan[1], Rhucha Joshi[1], Louis Ehwerhemuepha[1], and Michel Mikhael[1]

[1]Children's Hospital of Orange County

August 24, 2024

Abstract

Background: Preterm neonates with respiratory distress syndrome (RDS) who fail continuous positive airway pressure (CPAP) ventilation have higher risks for increased morbidity and mortality. **Objective:** To assess if machine learning, on multicenter data, may predict CPAP failure in preterm infants with RDS and allow proactive intervention to minimize CPAP failure burden and improve clinical outcomes. **Methods:** This study was conducted using the Oracle EHR Real-World Data (OERWD) database including preterm NICU admits between 2002-2023. CPAP failure was defined as the need for invasive mechanical ventilation within 72 hours of life. Demographics, admit vital signs, and laboratory values were retrieved to develop an explainable machine learning model using extreme gradient boosting (XGBoost). **Results:** 24,127 neonates from 27 NICUs qualified for the study with CPAP failure rate of 64.1%. FiO2 was the strongest predictor of CPAP failure followed by systolic blood pressure, temperature, birthweight, PaO2, oxygen saturation, heart rate, and gestational age followed in importance. Resulting XGBoost model attained an area under the receiver operator characteristic curve of 0.91 (95% CI: 0.90, 0.92) and an F-1 score of 0.87. **Conclusions:** CPAP failure can be predicted with high accuracy at admission to the NICU creating opportunities for early intervention and prevention of RDS related complications.

Early Prediction of CPAP Failure in Very Preterm Neonates Utilizing Machine Learning

Quinn Gates, MS*[1], Shruthi Janardhan, MD[1,2], Rhucha Joshi, MD[1,2], Louis Ehwerhemuepha, PhD[1], Michel Mikhael, MD[1,2]

[1]CHOC Children's Hospital, Orange, CA,[2]University of California Irvine, Irvine, CA

Corresponding Author Information

Name: Quinn Gates

Email: qgates1@choc.org

Phone Number: (949)981-4963

Address: 26662 Cadenas, Mission Viejo, CA 92691

Category of Study

Population Study

Impact

- CPAP Failure in preterm infants experiencing respiratory distress syndrome can be predicted with high degree of accuracy.

Posted on 24 Aug 2024 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.172448181.17701053/v1 | This is a preprint and has not been peer-reviewed. Data may be preliminary.

- Accurate prediction of CPAP failure will allow clinicians to make proactive interventions and reduce unnecessary intubation.
- To the best of our knowledge, this is the first study incorporating machine learning techniques to predict CPAP Failure on a multicenter scale.

Abstract

Background: Preterm neonates with respiratory distress syndrome (RDS) who fail continuous positive airway pressure (CPAP) ventilation have higher risks for increased morbidity and mortality.

Objective: To assess if machine learning, on multicenter data, may predict CPAP failure in preterm infants with RDS and allow proactive intervention to minimize CPAP failure burden and improve clinical outcomes.

Methods: This study was conducted using the Oracle EHR Real-World Data (OERWD) database including preterm NICU admits between 2002-2023. CPAP failure was defined as the need for invasive mechanical ventilation within 72 hours of life. Demographics, admit vital signs, and laboratory values were retrieved to develop an explainable machine learning model using extreme gradient boosting (XGBoost).

Results: 24,127 neonates from 27 NICUs qualified for the study with CPAP failure rate of 64.1%. FiO2 was the strongest predictor of CPAP failure followed by systolic blood pressure, temperature, birthweight, PaO2, oxygen saturation, heart rate, and gestational age followed in importance. Resulting XGBoost model attained an area under the receiver operator characteristic curve of 0.91 (95% CI: 0.90, 0.92) and an F-1 score of 0.87.

Conclusions: CPAP failure can be predicted with high accuracy at admission to the NICU creating opportunities for early intervention and prevention of RDS related complications.

Introduction

Respiratory distress syndrome (RDS) is a common cause of respiratory morbidity in premature infants due structural and functional immaturity of the lungs together with deficiency of alveolar surfactant, resulting in microatelectasis and low lung volumes [1,2]. The mainstay in the management of RDS is appropriate respiratory support and surfactant replacement as needed to achieve and maintain functional residual capacity (FRC) for better gaseous exchange.

Bronchopulmonary dysplasia (BPD) is the most common complication of prematurity, leading to many long-term morbidities and adverse neurodevelopmental outcomes [3]. The Neonatal Research Network collected data on over 34,000 infants born at 22–28 weeks' gestation between 1993–2012 and demonstrated significant increases in survival among infants born at 23- 25 weeks gestational age[4]. These infants are at a high risk of developing BPD – with an incidence of 60–80% [5]. Despite the increased use of antenatal corticosteroids, surfactant, and improved ventilation techniques, the incidence of BPD remains unchanged and may be increasing [6,7]. Several treatment strategies remain in place to prevent BPD, with one such strategy being avoidance of mechanical ventilation altogether to prevent ventilator-induced lung injury and use of early noninvasive continuous positive airway pressure (CPAP). Over the past decade, trials have demonstrated the feasibility and apparent benefits of providing early CPAP, including a reduced need for intubation and surfactant therapy [6,8,9]. The clinical outcomes for infants who succeed on CPAP are excellent, with low rates of BPD, mortality, intraventricular hemorrhage, retinopathy of prematurity, and lower risk for adverse neurodevelopmental sequelae in school-age [10].

However, despite best efforts, CPAP failure rates remain high and is associated with mortality and significant morbidity such as pneumothorax and BPD [5,11]. Identification of these infants in attempt to predict failure will allow early proactive intervention in the course of the disease, in turn reducing the complications of RDS and CPAP failure.

Previous research studies have evaluated variables that might predict CPAP failure. Several factors to predict treatment failure have been studied including demographics, perinatal and neonatal care variables, and was shown that increased fraction of inspired oxygen (FiO2) requirement in the first few hours of life (HOL) was

Posted on 24 Aug 2024 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.172448181.17701053/v1 | This is a preprint and has not been peer-reviewed. Data may be preliminary.

most consistent to predict CPAP failure [11–13]. These studies focused on perinatal and neonatal variables, known to be risk factors of severe RDS.

In this study, we assess whether the use of machine learning and electronic decision support systems may prove valuable to predict CPAP failure in preterm infants with RDS and allow early proactive intervention to minimize CPAP failure burden.

Methods

This retrospective cohort study was approved by the Children's Hospital of Orange County Institutional Review Board (IRB #2008107).

Data Sources, Patients, and Variables

Oracle EHR Real-World Data (OERWD)– a large multicenter electronic health records (EHR) database – was used for this study. The database contains data from more than 125 US health systems as of September 2023. OERWD is fully de-identified, encrypted, and secured in compliance with the Health Insurance Portability and Accountability Act of 1996 privacy regulation [14,15]. Details about the database are available in the data descriptor paper by Ehwerhemuepha et. al. 2022[14].

Cohort for the study was defined and retrieved from the database as preterm infants with gestational age of 32 weeks or less and/or birthweight less than 1500 gram. Patients' assigned sex, race, ethnicity, healthcare plan, gestational age, birthweight, vital signs, FiO2, partial pressure of oxygen (PaO2), and partial pressure of carbon dioxide (PCO2) values. Perinatal or maternal information was not retrieved due to inability to link mother and infant data in the database used. CPAP failure was defined as the introduction of invasive mechanical ventilation within 72 hours of birth.

Machine Learning Modeling

Data collected was split into training (75%) and test (25%) sets. Extreme gradient boosting (XGBoost) - an implementation of stochastic gradient boosting – was selected given its ability to capture complex nonlinear relationships between outcome and predictor variables in the presence of missing data [16,17]. Ten-fold cross-validation on the training set was used to determine the optimal hyperparameters of the model from a grid of values consisting of learning rates (to control model convergence); and maximum tree depth (2, 4, or 6) to control complexities of the trees built while setting other hyperparameters to their default values. The optimal hyperparameters were used to develop a final model on the training dataset. Variables used in the model were ranked for importance using the "Gain" metric, which measures the improvement in model performance by a feature on given branches of the trees it is on. The Shapley Additive Explanation (SHAP) values were used to provide visual explanation of the risk of CPAP failure given specified values of a feature. Higher SHAP values imply greater risk of CPAP failure. The test set was used to evaluate unbiased estimates of model performances such as the area under the receiver operator characteristic curve (AUROC), area under precision-recall curve (AUCPR), F-1 score, positive and negative predictive values, and relative risk of failing CPAP given a positive prediction.

This study was carried out using the Oracle Health HealtheDataLab platform as well as the R Statistical Programming Language.

Results

There were 24,127 preterm infants who qualified for inclusion into the cohort from 27 neonatal intensive care units in the US. This comprised of infants born between February 2002 and March 2023. The cohort consisted of 15,457 (64.1%) infants who failed CPAP. Among the infants who failed CPAP, there were 2% American Indian or Alaskan, 2.5% Asian, 15.8% Black or African American, 0.2% Native Hawaiian or Other Pacific Islander, 63.6% White or Caucasian, 2.2% with multiple races, and 13.7% other or unknown race. The ethnic distribution was 30.6% infants of Hispanic or Latino origin with 14.4% of unknown ethnic group. Of those in the cohort, 47.3% were assigned female at birth (52.7% as males); 50.4% were on governmental

Posted on 24 Aug 2024 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.172448181.17701053/v1 | This is a preprint and has not been peer-reviewed. Data may be preliminary.

insurance plans; 28.6% on Commercial/private insurance; 1.8% on Self-pay; and 19.2% with other or unknown insurance. Please refer to Table 1 for summary statistics on the entire cohort.

Splitting the data into a training (75%) and a test (25%) set results in data from 18,095 patients in the training set and 6,032 in the test set. Please refer to Table 2 on the Bivariate statistics by CPAP failure status on the data used for training the model. Hyperparameter tuning resulted in the selection of an XGBoost model with 64 trees, maximum tree depth of 6, and learning rate of 0.1. Feature importance is shown in Figure 1 indicating that the strongest predictors of CPAP failure as FiO2, systolic blood pressure, body temperature, PaO2, birthweight, oxygen saturation, gestational age, and heart rate.

The SHAP values indicate that there is increased risk for CPAP failure as FiO2 increases up to 32%. This risk was sustained for higher values of FiO2. In a similar way, the SHAP values indicate an increase in risk with fever or higher body temperature. Values of birth weight between 500 grams and 1000 grams exposed patients to higher risk of failure compared to those with birthweight between 1200 and 2000 grams. Most patients had oxygen saturation between 60 and 100% such that risk for failure reduced with increase in oxygen saturation levels. Similarly, the risk for failure decreased with increase gestational age. SHAP values for heart rate indicate a non-linear relationship such that lower-than-normal heart rates were associated with higher risk of failure. Values between 30 to 60 mmHg for pCO2 were associated with lower risk of failure compared to higher values. Higher respiratory rates were associated with reduced risk of failure. And infants assigned male at birth were at higher risk of failing CPAP. The SHAP values of the top 12 features (excluding insurance/payer) are shown in Figure 2.

Model performance on the independent test set using the area under the receiver operator characteristic curve was 0.91 (95% CI: 0.90, 0.92). Balancing sensitivity and specificity (in determining a classification threshold) results in sensitivity of 0.86 (95% CI: 0.85, 0.87); specificity of 0.82 (95% CI: 0.80, 0.83); positive predictive value of 0.89 (95% CI: 0.88, 0.90); negative predictive value of 0.76 (95% CI: 0.75, 0.78); F1 score of 0.87; predicted probability threshold for classification at 0.69; and 10 of 11 predictions were true positive predictions in the test set. Area under the precision-recall curve was 0.93. A visual representation of the areas under the curves is shown in Figure 3.

Discussion

Neonates who fail CPAP are at increased risk of pneumothorax, prolonged respiratory support, bronchopulmonary dysplasia, and death[18]. Identifying these neonates shortly after birth is crucial for implementing targeted strategies to reduce such complications.

In our study, we created a machine learning model to predict the risk of CPAP failure in preterm neonates born less 32 weeks gestation and or with birth weight less than 1500 grams. To our knowledge, this is the first study describing the utility of machine learning in prediction of CPAP failure using large multicenter data. CPAP failure was defined as the need for invasive mechanical ventilation within the first 72 hours of life.

The rate of CPAP failure in our cohort was 64.1%. This is higher than previously reported rates of 21% to 51% [10,19–22]. The wide range of these rates is likely related to institutional differences in initial respiratory management and including older cohort when practices such as routine intubation of infants of certain gestational ages were common [21]. Since our cohort was comprised of infants from 27 neonatal intensive care units, we could not account for variability between the centers in criteria for invasive ventilation, and surfactant administration practices (including indication and mode of surfactant delivery).

The strongest predictor of CPAP failure in our model was FiO2. In our model, the highest risk of CPAP failure was at a FiO2 of 0.31. This is consistent with prior studies that reported FiO2 in the first few hours of life as predictive of CPAP failure, with FiO2 cutoffs ranging from 0.25 to 0.3 [10,19,20,23].

We found an increasing risk of CPAP failure with elevated body temperature. Other variables that had a high variable importance ranking were systolic blood pressure and arterial PaO2. Several studies report birth weight as a significant predictor for CPAP failure[10,23,24], and it was one of the top 5 variables that

Posted on 24 Aug 2024 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.172448181.17701053/v1 | This is a preprint and has not been peer-reviewed. Data may be preliminary.

contributed to our prediction model. Neonates with a birth weight of 500-1000 grams had a higher risk of failing CPAP compared to those with a birth weight of 1200-2000 grams. We also found that the risk of CPAP failure decreased with increasing gestational age as expected.

Our study is unique in that it provides a machine learning model for predicting CPAP failure using multi-center data. The benefit of machine learning is that it accounts for complex relationships between variables and does not assume that these predictors are independent of each other. It can use large data sets, such as our cohort, to generate a model that is validated as the algorithm is created [25]. We tested our model with a subset of our cohort and found that area under the receiver operator characteristic curve was 0.91 and area under the precision-recall curve was 0.93, indicating very high model performance.

There were several limitations in this study including the absence of a standard definition of CPAP failure. Consequently, every multicenter study may be limited by institutional differences in classifying patients who failed CPAP and corresponding treatment recommendation.

The objective of this study is development of the earliest warning system predicting risk of CPAP failure by using the very first set of clinical data as predictors. While this is very helpful, there may be increased accuracy but reduced time for intervention for predicting at later times into the course of the NICU stay. Future multicenter studies may provide more light into the effectiveness of longitudinal follow-up till CPAP failure or discharge otherwise.

There were missing data which we accounted for in the choice of machine learning algorithm. However, although these algorithms can capture relationships between input variables in the presence of missing data, the accuracy of the model may have been impacted. Due to the de identification of data in OERWD, a link between neonatal charts with maternal charts was not possible. Therefore, the model does not include important perinatal factors, such as antenatal steroids exposure, mode of delivery and Apgar scores.

In conclusion, we describe a machine learning model to predict CPAP failure in preterm neonates with RDS. The most significant predictor in our model was FiO2, where we observed a direct relationship between FiO2 and CPAP failure up until a value of 0.32 after which the chances of CPAP failure remained relatively stable (see Figure 2). Other predictors of importance were gestational age, birth weight and first PCO2 value. We tested the model's ability to predict CPAP failure in a subset of our cohort and found high performance. Implementation of this model into the electronic medical record can facilitate early prediction of CPAP failure and identify neonates who may benefit from targeted interventions to maximize the success of non-invasive ventilation.

References

1. Reuter, S., Moser, C. & Baack, M. Respiratory Distress in the Newborn. *Pediatr Rev* **35** , 417–429 (2014).

2. Yadav, S., Lee, B. & Kamity, R. *Neonatal Respiratory Distress Syndrome* . (2023).

3. Jensen, E. A. Prevention of Bronchopulmonary Dysplasia: A Summary of Evidence-Based Strategies. *Neoreviews* **20** , e189–e201 (2019).

4. Stoll, B. J. *et al.* Trends in Care Practices, Morbidity, and Mortality of Extremely Preterm Neonates, 1993-2012. *JAMA* **314** , 1039–51 (2015).

5. Wright, C. J., Sherlock, L. G., Sahni, R. & Polin, R. A. Preventing Continuous Positive Airway Pressure Failure: Evidence-Based and Physiologically Sound Practices from Delivery Room to the Neonatal Intensive Care Unit. *Clin Perinatol* **45** , 257–271 (2018).

6. Dunn, M. S. *et al.* Randomized Trial Comparing 3 Approaches to the Initial Respiratory Management of Preterm Neonates.*Pediatrics* **128** , e1069–e1076 (2011).

7. Morley, C. J. *et al.* Nasal CPAP or Intubation at Birth for Very Preterm Infants. *New England Journal of Medicine* **358** , 700–708 (2008).

Posted on 24 Aug 2024 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.172448181.17701053/v1 | This is a preprint and has not been peer-reviewed. Data may be preliminary.

8. SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network *et al.* Early CPAP versus surfactant in extremely preterm infants. *N Engl J Med* **362** , 1970–9 (2010).

9. Avery, M. E. *et al.* Is Chronic Lung Disease in Low Birth Weight Infants Preventable? A Survey of Eight Centers. *Pediatrics* **79** , 26–30 (1987).

10. Gulczyńska, E., Szczapa, T., Hożejowski, R., Borszewska-Kornacka, M. K. & Rutkowska, M. Fraction of Inspired Oxygen as a Predictor of CPAP Failure in Preterm Infants with Respiratory Distress Syndrome: A Prospective Multicenter Study. *Neonatology* **116** , 171–178 (2019).

11. Abdallah, Y. *et al.* CPAP failure in the management of preterm neonates with respiratory distress syndrome where surfactant is scarce. A prospective observational study. *BMC Pediatr* **23** , 211 (2023).

12. De Jaegere, A. P., van der Lee, J. H., Canté, C. & van Kaam, A. H. Early prediction of nasal continuous positive airway pressure failure in preterm infants less than 30 weeks gestation. *Acta Paediatr* **101** , 374–379 (2012).

13. Dargaville, P. A. *et al.* Continuous Positive Airway Pressure Failure in Preterm Infants: Incidence, Predictors and Consequences.*Neonatology* **104** , 8–14 (2013).

14. Ehwerhemuepha, L. *et al.* Cerner real-world data (CRWD) - A de-identified multicenter electronic health records database. *Data Brief* **42** , 108120 (2022).

15. Ehwerhemuepha, L. *et al.* HealtheDataLab - a cloud computing solution for data science and advanced analytics in healthcare with application to predicting multi-center pediatric readmissions. *BMC Med Inform Decis Mak* **20** , 115 (2020).

16. Chen, T., He, T. & Benesty, M. Xgboost: extreme gradient boosting.*R package version 0.4-3* 1–4 (2015).

17. Chen, T. *et al.* xgboost: Extreme Gradient Boosting. Preprint at (2019).

18. Dargaville, P. A. *et al.* Incidence and Outcome of CPAP Failure in Preterm Infants. *Pediatrics* **138** , (2016).

19. Kakkilaya, V. *et al.* Early predictors of continuous positive airway pressure failure in preterm neonates. *Journal of Perinatology* **39** , 1081–1088 (2019).

20. Dargaville, P. A. *et al.* Continuous Positive Airway Pressure Failure in Preterm Infants: Incidence, Predictors and Consequences.*Neonatology* **104** , 8–14 (2013).

21. Rocha, G. *et al.* Failure of early nasal continuous positive airway pressure in preterm infants of 26 to 30 weeks gestation.*Journal of Perinatology* **33** , 297–301 (2013).

22. Fuchs, H., Lindner, W., Leiprecht, A., Mendler, M. R. & Hummler, H. D. Predictors of early nasal CPAP failure and effects of various intubation criteria on the rate of mechanical ventilation in preterm infants of <29 weeks gestational age. *Arch Dis Child Fetal Neonatal Ed* **96** , F343-7 (2011).

23. De Jaegere, A. P., van der Lee, J. H., Cante, C. & van Kaam, A. H. Early prediction of nasal continuous positive airway pressure failure in preterm infants less than 30 weeks gestation. *Acta Paediatr* **101** , 374–379 (2012).

24. Ammari, A. *et al.* Variables Associated with the Early Failure of Nasal CPAP in Very Low Birth Weight Infants. *J Pediatr* **147** , 341–347 (2005).

25. Sidey-Gibbons, J. A. M. & Sidey-Gibbons, C. J. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol* **19** , 64 (2019).

Funding Support Statement

This work was supported by Chiesi Farmaceutici S.p.A.

Posted on 24 Aug 2024 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.172448181.17701053/v1 | This is a preprint and has not been peer-reviewed. Data may be preliminary.

Competing Interests

The authors declare no competing interests.

Ethics Approval and Consent to Participate

No patient consent was applicable for this paper.

Main Figure Legend

Title	Description
Figure 1	Feature importance (Gain) ranking for predicting CPAP failure
Figure 2	Shapley additive explanations of risk for CPAP failure by first measured vital signs and selected laboratory test
Figure 3	Areas under the receiver-operator and precision-recall curves
Table 1	Summary statistics on all variables recovered
Table 2	Bivariate analyses of variables

Hosted file

`Figure 1.docx` available at https://authorea.com/users/817629/articles/1217858-early-prediction-of-cpap-failure-in-very-preterm-neonates-utilizing-machine-learning

Hosted file

`Figure 2.docx` available at https://authorea.com/users/817629/articles/1217858-early-prediction-of-cpap-failure-in-very-preterm-neonates-utilizing-machine-learning

Hosted file

`Figure 3.docx` available at https://authorea.com/users/817629/articles/1217858-early-prediction-of-cpap-failure-in-very-preterm-neonates-utilizing-machine-learning

Hosted file

`Table 1.docx` available at https://authorea.com/users/817629/articles/1217858-early-prediction-of-cpap-failure-in-very-preterm-neonates-utilizing-machine-learning

Hosted file

`Table 2.docx` available at https://authorea.com/users/817629/articles/1217858-early-prediction-of-cpap-failure-in-very-preterm-neonates-utilizing-machine-learning

Hosted file

`Tables.docx` available at https://authorea.com/users/817629/articles/1217858-early-prediction-of-cpap-failure-in-very-preterm-neonates-utilizing-machine-learning