PREDICTION OF CPAP FAILURE IN THE NEONATAL ICU

Laura Livant, MS, Colleen M. Ennett, PhD,
Philips Research North America, Briarcliff Manor, NY

ABSTRACT

Continuous Positive Airway Pressure (CPAP) is now the preferred treatment for neonatal patients suffering from less severe respiratory distress upon admission to the Neonatal Intensive Care Unit (NICU). If a patient "fails" CPAP, meaning the patient does not respond to CPAP therapy and needs to be transitioned to invasive mechanical ventilation, this could cause later complications for the infant's development. This study suggests predictors for CPAP failure based on clinical data from the MIMIC II (Multi-parameter Intelligent Monitoring for Intensive Care, Version 2) database. A reference data set was created by selecting neonatal patients who were <36 weeks gestational age, born with Respiratory Distress Syndrome, placed on CPAP within 24 hours of NICU admission, and on CPAP for >4 hours. From the 22 features considered as possible predictors of CPAP failure, the final features were chosen based on a minimum number of patients who had the feature recorded, a statistically significant difference of that feature between the pass/fail groups of patients using F-score, and consultation with a neonatologist. Logistic regression with 100X bootstrapping was used to assess combinations of one, two and three features with and without interaction terms. The set of three features that showed both the highest AUC (Area Under ROC Curve) and highest sensitivity at a specificity of 90% were Oxygen Saturation (SaO$_2$), Total Cycle Time, and Indirect Bilirubin. This set of three features resulted in an AUC of 96.5%, Sensitivity of 88.2%, and Specificity of 90%.

INTRODUCTION

In the last decade, Continuous Positive Airway Pressure (CPAP) has increasingly become the treatment of choice for neonatal patients in respiratory distress upon admission to the Neonatal Intensive Care Unit (NICU), for patients where invasive mechanical ventilation might be avoided. While the indications for initiating CPAP therapy are somewhat well-established, there are very few clinically-based studies or guidelines that exist to advise when to transition a patient off of CPAP [1]. Todd et al. [2] determined that stopping CPAP abruptly results in better outcomes than gradual weaning, however, only 32% of infants in their study passed CPAP on the first attempt. Todd et al. acknowledge that "when to wean" is still a contentious issue, and is usually done *ad hoc*. Patients who are initially placed on CPAP but do not respond and have to subsequently be placed on invasive ventilation suffer greater distress and are at higher risk of complications due to the delay of needed ventilation and in many cases due to the delay of surfactant delivery [3]. Putting a patient on invasive ventilation unnecessarily introduces risk for complication, tissue damage and various ventilator-related injuries [4].

Because neonatal patients vary greatly in their level of physiological and neurological development, the problem of how to treat respiratory distress in premature infants is a difficult one. In addition, many factors such as birth weight, health of the mother, and other environmental and genetic factors may play a role in a premature newborn's respiratory condition. An intelligent model that can identify combinations of indicators that predict when a neonatal patient should be placed on invasive ventilation could aid the clinician in choosing treatment options that reduce the prevalence of "CPAP Failure", when a patient placed on CPAP must be transitioned to invasive ventilation. "Passing CPAP" refers to a patient being successfully weaned from CPAP ventilation to less invasive ventilation support or no ventilation support. Currently the decisions of when and how to transition patients off of CPAP

are made largely based on clinician training and experience, and on guidelines (if they exist) from the hospital or region [1],[2]. Algorithms that provide guidance derived from clinically-based prediction models could inform clinicians in their decision-making during the critical first hours and days of treating a premature infant.

There are a great deal of collected clinical data available that may include valuable information that could aid clinicians in deciding when to provide ventilation and which type of ventilation has the best predictive outcome for a particular patient. This data analysis study took an existing clinical database and extracted indicators that identified which neonatal patients failed CPAP and which passed CPAP. This study was conducted to determine which features may be used to predict CPAP failure using the MIMIC II (Multi-parameter Intelligent Monitoring for Intensive Care, Version 2) patient database [5].

METHODS

Clinical Database

Clinical data were drawn from the MIMIC II Database, Version 2.6, which contains clinical patient data for approximately 32,000 patients, where 7839 are NICU patients. The data was collected from patients at Beth Israel Deaconess Medical Center in partnership with Philips Healthcare and Massachusetts Institute of Technology, funded by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and the National Institutes of Health (NIH). The data collection study was conducted between 2001 and 2008.

The original MIMIC II database, accessed through a MySQL application, was imported into a MATLAB database for further analysis. Features considered to be relevant to this study based on the literature review and advice from clinical experts were extracted from the original database and additional features specific to ventilation were calculated from the existing data.

The patient data were collected for general purposes, not specifically to study ventilation needs. Therefore, there are some issues with using these data to analyze neonatal ventilation in a meaningful way. The frequency of data collected, gaps in data, conflicting data recorded, and lack of certain data specific to known neonatal ventilation conditions are examples of the issues in using the MIMIC II database. In addition, the MIMIC-II database does not provide clear indication for when a patient was placed on or off of a particular type of ventilation, nor does it clearly indicate the specific type of ventilation. There are several ventilator-related chart events that when evaluated together give a fairly good record of the ventilation status. A weighting-based approach was developed that assigns a score for each ventilation period based on a combination of ventilation-related feature settings. This score was used to determine what type of ventilation the patient was on, and when the patient was put on and taken off of that type of ventilation [6]. Ventilator and CPAP on/off times were manually validated by checking the nursing notes associated with each patient included in the study. Patients were excluded from analysis if the ventilator and CPAP on/off times could not be verified.

Patient Selection

A patient data set was constructed by extracting neonatal patients that were placed on CPAP therapy as their first mode of ventilation support. Inclusion criteria were Respiratory Distress Syndrome (RDS) diagnosis as identified by the ICD-9 code, <36 weeks gestational age at birth, and CPAP therapy within 24 hours of NICU admission. While having an adequate number of patients is important, we also required a sufficiently long period of data collected to be able to assess the success or failure of CPAP therapy. This created a trade-off in that for the patient data collected, the neonates who failed CPAP generally did so very quickly, whereas the patients who passed CPAP were on CPAP for significantly longer durations on average. The average CPAP ventilation duration of patients who failed CPAP was 0.69 days, whereas the average CPAP duration of patients who passed CPAP was 8.5 days. Figure 1 shows the distribution of the number of patients who pass and fail CPAP and are still on CPAP support after the durations specified. The large drop in patients that fail CPAP after the first four hours shows that a significant number of patients were transitioned
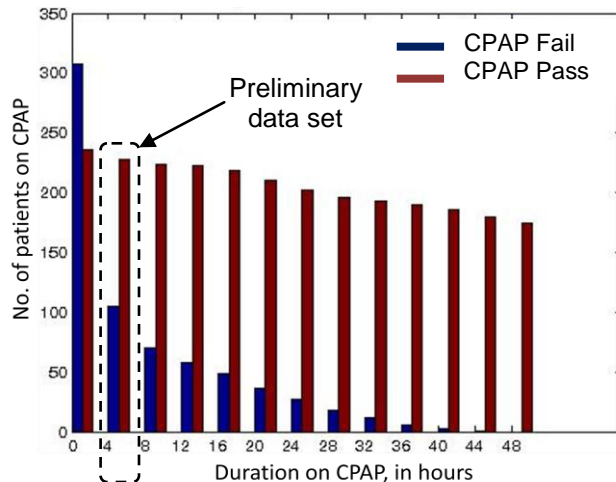
Figure 1. Number of patients on ventilation after duration who will Fail and Pass CPAP

to invasive ventilation within the first four hours of NICU admission.

To keep enough patients who failed CPAP in the study while including a sufficient duration of ventilation support to use reliable data, the patient data set was limited to patients who were on CPAP for at least four hours. Patients who had zero or erroneous durations as determined by the database timestamps were excluded, as well as patients who died while being treated in the NICU. The final Patient Data Set using the above criteria consisted of 308 patients; 94 patients failed CPAP and 214 patients passed CPAP.

Feature Selection

The original feature set evaluated had 122 features. These features included chart events, lab events, and events recorded from nursing notes. In addition, several relevant ventilation-specific features were calculated from existing features, such as PF Ratio, calculated as PaO_2/FiO_2, or partial pressure of oxygen divided by fraction of inspired oxygen. From this initial set of features, a final set was extracted using a combination of methods. Features were evaluated against the final patient set for two-, four-, and six-hour periods prior to CPAP ventilation transition for each of the patients. Features that were measured for less than 50 of the patients in the final patient set for the period prior to transition were excluded. In addition, features were evaluated based on how statistically distinct the pass/fail sets were for

that feature, as determined by Youden's Index, and confirmed by F-score. Features with Youden's index value of less than 0.25 were excluded from the study. In addition, Pearson's correlation coefficient, ρ, was calculated for each feature against every other feature. For features that were correlated with ρ values greater than 0.7, the feature that had the lower F-score was excluded.

After evaluating the features statistically, we evaluated the features based on prior literature review and consulted a neonatologist to finalize the feature list. We included Birth Weight and Gestational Age, which would have otherwise been discounted according to the statistical evidence, due to their prevalence in literature [7],[8]. We excluded Mean Airway Pressure (MAP) due to the unreliability of this feature for CPAP patients. We also discounted Fraction of Inspired Oxygen (FiO_2) as a potential feature, although in initial analysis this feature showed up as a very strong predictor of CPAP failure, and it is identified as a strong indicator of CPAP failure in the literature. After looking at a breakdown of FiO_2 in the neonatal patients in this study, we found that most neonatal patients failed CPAP with FiO_2 near 0.6. While we do not have insight into the practices of the hospital that recorded these data, it seems that they transitioned patients to invasive ventilation at an FiO_2 setting between 0.4 and 0.6. Thus the discriminating power of this feature reflects the clinicians' practices rather than a predictor of CPAP failure.

After applying all of these selection criteria, 22 of the 122 original features were extracted and used for the data analysis and evaluation.

Regression Analysis

To identify a list of features that could be used to predict CPAP failure, logistic regression was performed along with 100-times non-parametric bootstrapping, to account for features with non-normal distributions. Combinations of one, two and three features of the 22 selected features were evaluated with and without interaction terms. Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) and Sensitivity were evaluated at 90% Specificity, because high Specificity is important for clinical applications to reduce

false positives which would unnecessarily alert the clinician for infants who are not at risk.

RESULTS

Using the patient data set of 308 patients, and the final feature list of 22 features, regression analysis was performed for combinations of one, two and three features, with and without interaction terms. The best results, defined as highest AUC and Sensitivity with Specificity of 90%, were found with the combination of three features: Oxygen Saturation (SaO_2), Total Cycle Time (TT, calculated as 60/Respiration Rate), and Indirect Bilirubin (IB) with Interaction Terms (IT), where the interaction terms are the combination of the three features in the algorithm (SaO_2*TT*IB). This resulted in an AUC of 96.5% and a Sensitivity of 88.2% at a Specificity of 90%. The logistic regression bootstrap model coefficients are estimated in the Equation (1). The ROC curve for this model for the training data and bootstrap model is shown in Figure 2.

$$z = 62 - 0.50*SaO_2 - 19*TT - 1.1*IB + 0.1*IT \qquad (1)$$

DISCUSSION

We developed a new model for identifying predictive features for CPAP failure. This model used the MIMIC-II neonatal patient dataset to discover a combination of features that provide promising results in predicting CPAP failure in

Figure 2. ROC Curve for SaO_2+TT+IB+IT

premature infants with respiratory distress. The features discovered are routinely measured and recorded for neonatal patients.

Out of the features identified in the model, Oxygen Saturation, or the measure of oxygen in the blood, has been mentioned in literature as an indicator of CPAP failure [9], and is a logical indicator of a failure to respond to CPAP ventilation. Likewise, Total Cycle Time, or the measure of frequency of breathing, is also a measurement one might expect would indicate a failure to respond to CPAP ventilation. Indirect Bilirubin, however, was initially an unexpected result. Indirect or unconjugated bilirubin is a common indicator of jaundice in infants. It has also been shown to be an anti-oxidant, especially in newborns. Upon further study, it was found that when premature infants are exposed to oxidative stress, moderate levels of unconjugated bilirubin may be generated as a protective factor [10]. This is an interesting finding and may explain why Indirect Bilirubin was found as a predictor of CPAP failure in combination with Oxygen Saturation and Total Cycle Time.

There are several limitations to this model. As mentioned earlier, the lack of frequency of data collected and the lack of other features recorded that have been known to indicate respiratory distress in infants (e.g., antenatal steroids taken by the mother) could be hiding other features that influence CPAP failure in neonates. In addition, feature selection was determined by evaluating each feature individually, whereas the model accounted for combinations of features. Also, bootstrapping was employed in model building but not in feature selection. Feature selection and model building should be done using the same techniques and criteria. The results from this model should be evaluated in future studies to confirm and refine the findings here.

This type of model could further develop ventilation patient monitoring capabilities to aid clinical decision support. Establishing prediction criteria for the timing of CPAP transition based on computational analysis of clinical data can support clinicians in making time-critical decisions, which could lower the occurrence of ventilation-related injury and complications among premature infants in the NICU.
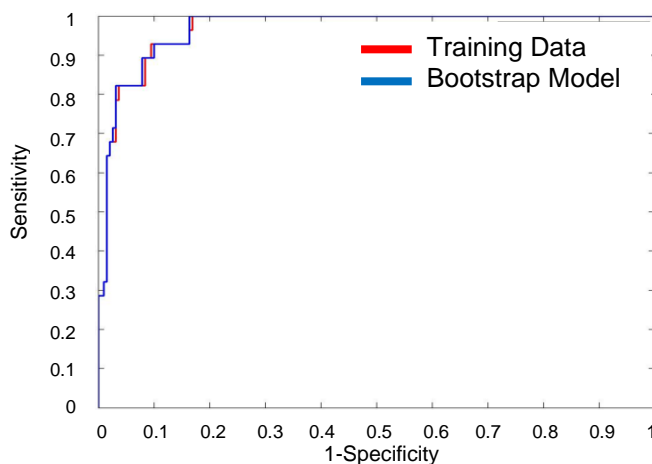
REFERENCES

[1] C. Roehr, G. Schmalisch, A. Khakban, H. Proquitte, R. Wauer, "Use of continuous positive airway pressure (CPAP) in neonatal units – a survey of current preferences and practice in Germany," *European Journal of Medical Research*, pp. 139-44, 2007.

[2] D. A. Todd, A. Wright, M. Broom, M. Chauhan, S. Meskell, et al., "Methods of weaning preterm babies <30 weeks gestation off CPAP: a multicentre randomised controlled trial," Arch. Dis. Child Fetal Neonatal Ed., vol. 97, no. 4, July 2012.

[3] H. Fuchs, W. Lindner, A. Leiprecht, M. R. Mendler, H. D. Hummler, "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*, pp. F343-7, 2012.

[4] J. L. Wei, J. Bond, "Management and prevention of endotracheal intubation injury in neonates," Curr. Opin. Otolaryngol. Head Neck Surg., vol. 19, no. 6, pp. 474-477, Dec. 2011.

[5] M. Saeed, M. Villarroel, A. T. Reisner, et al., "Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): A public-access intensive care unit database," Crit. Care Med., vol. 39, no. 5, pp. 952–960, May 2011.

[6] A. Mikhno and C. M. Ennett, "Prediction of extubation failure for neonates with respiratory distress syndrome using the MIMIC-II Clinical Database," *Proceedings of 34th Annual International Conference of the IEEE EMBS,* pp. 5094-5097, 2012.

[7] A. Ammari, S. Mandhir, V. Milisavljevic, R. Sahni, D. Bateman, U. Sanocka, "Variables Associated with the Early Failure of Nasal CPAP in Very Low Birth Weight Infants," *Newborn and Infant Nursing Reviews*, pp. 68-75, 2006.

[8] M. S. Pillai, M. J. Sankar, K. Mani, R. Agarwal, V. K. Paul, A. K. Deorari, "Clinical Prediction Score for Nasal CPAP Failure in Pre-Term VLBW Neonates with Early Onset Respiratory Distress," *Journal of Tropical Pediatrics*, pp. 274-9, 2011.

[9] A. P. De Jaegere, J. H. van der Lee, C. Canté, A. H. van Kaam, "Early prediction of nasal continuous positive airway pressure failure in preterm infants less than 30 weeks gestation," Acta Paediatrica, vol. 101, pp. 374-379, 2012.

[10] L. Thrall, "The Physiologic Role of Bilirubin in the Newborn," *The Matrona Research and Literature,* http://www.thematrona.com/apps/blog/the-physiologic-role-of-bilirubin-in-the, 2012.