Novel pediatric-automated respiratory score using physiologic data and machine learning in asthma

Amanda I. Messinger MD | Nam Bui | Brandie D. Wagner PhD | Stanley J. Szefler MD | Tam Vu PhD | Robin R. Deterding MD

1Department of Pediatrics, Colorado School of Medicine, The Breathing Institute, University of Colorado, Children’s Hospital Colorado, Aurora, Colorado
2Department of Computer Science, University of Colorado Boulder, Boulder, Colorado
3Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado, Aurora, Colorado

Correspondence
Amanda I. Messinger, Department of Pediatrics, Colorado School of Medicine, The Breathing Institute, University of Colorado, Children’s Hospital Colorado, 13123 East 16th Avenue, Box 395, Aurora, CO 80045.
Email: amanda.messinger@childrenscolorado.org

Abstract
Objectives: Manual clinical scoring systems are the current standard used for acute asthma clinical care pathways. No automated system exists that assesses disease severity, time course, and treatment impact in pediatric acute severe asthma exacerbations. Working hypothesis: machine learning applied to continuous vital sign data could provide a novel pediatric-automated asthma respiratory score (pARS) by using the manual pediatric asthma score (PAS) as the clinical care standard.

Methods: Continuous vital sign monitoring data (heart rate, respiratory rate, and pulse oximetry) were merged with the health record data including a provider-determined PAS in children between 2 and 18 years of age admitted to the pediatric intensive care unit (PICU) for status asthmaticus. A cascaded artificial neural network (ANN) was applied to create an automated respiratory score and validated by two approaches. The ANN was compared with the Normal and Poisson regression models.

Results: Out of an initial group of 186 patients, 128 patients met inclusion criteria. Merging physiologic data with clinical data yielded >37,000 data points for model training. The pARS score had good predictive accuracy, with 80% of the pARS values within ±2 points of the provider-determined PAS, especially over the mid-range of PASs (6-9). The Poisson and Normal distribution regressions yielded a smaller overall median absolute error.

Conclusions: The pARS reproduced the manually recorded PAS. Once validated and studied prospectively as a tool for research and for physician decision support, this methodology can be implemented in the PICU to objectively guide treatment decisions.

KEYWORDS
asthma, clinical prediction, machine learning

1 | INTRODUCTION

The use of pediatric clinical scoring systems in clinical care and research has exploded in the last decade. Clinical scoring systems appeal to multiple stakeholders because they are quantitative, can be validated and improve patient outcomes. Pediatric asthma is no exception; as the most common chronic disease of childhood, development of clinical scores and guidelines have helped to streamline and improve pediatric asthma care delivery. Many hospitals have developed clinical care guidelines for management of acute asthma exacerbations built around manual, provider-determined asthma severity scores. Examples include the pediatric asthma score (PAS) and pediatric respiratory assessment measure. These scores contain similar elements but are customized...
to individual hospitals. The use of these scores to guide inpatient treatment of acute asthma exacerbations has improved patient outcomes including reduced length of stay, decreased admission rates, and decreased medication burden in both the emergency department and inpatient wards.

Respiratory scores like PAS contain subjective elements like auscultation, have limited interrater reliability, and are less sensitive in older children and adolescents. They are used intermittently and depend on frequent reassessment, increasing the burden on staff. The provider-determined PAS score used in our institution includes qualitative measurements (Figure S1). Despite being useful on the wards, PAS scores are measured inconsistently in our intensive care unit (ICU) and not regularly used to make care decisions.

There is significant variation in care of pediatric intensive care unit (PICU) patients with asthma, both within and across centers and hospitals. Despite the existence of stepwise guidelines for outpatient asthma management, there are no national or international guidelines to direct severe acute asthma exacerbation management in the PICU. Validated scores for asthma severity can be useful in clinical and quality research on asthma therapies such as intravenous bronchodilators and use of noninvasive ventilation.

The objective of this study was to produce a novel, automated acute asthma severity score using machine learning. Machine learning is built on a foundation of mathematics, logic, probability, neuroscience, and decision theory. These foundational building blocks are used to generate computer algorithms that can keep a record of the relative strength of associations between data elements (similar to memory) through repeated training sessions. As a result, machine learning can identify patterns in complex data and then uses those patterns to construct models that can predict outcomes without relying on explicit human-generated programming code. Supervised learning is a subfield of machine learning based on tools of classification and regression; it depends on the input of a labeled training data set to help the computer “learn” relationships. That knowledge is then applied to an independent testing data set. The accuracy of outputs can then be analyzed. Artificial neural networks (ANNs) are a specific set of algorithms modeled after the structure of the human brain, designed to cluster and classify data and subsequently produce novel outputs.

We hypothesized that application of machine learning algorithms to passively collected vital sign data (heart rate, respiratory rate, and oxygen saturation) in critically ill pediatric asthma patients can generate a pediatric-automated asthma respiratory score (pARS) that could eventually replace PAS. Once created, the pARS can be validated and applied prospectively in the PICU, wards and emergency department to aid in clinical research and provider decision support without increasing the burden of staff or utilizing subjective measures.

2 | METHODS

This was a single center study conducted at a large quaternary children’s hospital. The Colorado Multiple Institutional Review Board (COMIRB 16-1359) approved this study. Eligible patients were identified during a 1-year period, 1 January 2016 to 1 January 2017 using data collected and stored during their clinical care.

2.1 | Patient selection

Inclusion criteria included patients admitted to the PICU age 2 to 18 years old with diagnosis codes for status asthmatics across all severities. Data from patients who had international classification of diseases (ICD) 9/ICD 10 procedure codes within the encounter for intubation, or received continuous invasive mechanical ventilation were excluded. Diagnosis codes for other potentially confounding chronic respiratory and neurologic conditions also disqualified patients (Table S1 and S2).

2.2 | Data collection

Demographic variables and time-stamped clinical data including charted PAS score, respiratory support, and medications were obtained from the electronic health record (EHR). Using bed numbers and admission/discharge time stamps, each patient’s continuous vital sign information was manually extracted from a central research database. This database stores vital sign data (numeric and waveform) from PICU patients attached to the Phillips monitors. Data extracted for this study included time-stamped values for heart rate, respiratory rate, and pulse oximetry.

2.3 | Data preparation

The EHR respiratory flow sheet data were aligned with vital sign data using date and time stamps for PAS score in Matlab (version R2017a). The recorded PAS was used as an outcome to train the supervised machine learning models. Patient data with complete records of three parameters (heart rate, respiratory rate, and pulse oximetry readings) overlapping with PAS time points were included. Patient data with incomplete alignment were excluded to create the final study cohort of 128 patients. For each PAS score recorded in the medical record, 20 minutes of vital sign measurements were associated with one PAS score via a standard one-to-many matching strategy to create discrete time periods for algorithm training (Figure S2). To control for age-based variability in heart rate and respiratory rate, z-scores for each patient’s heart rate and respiratory rate were calculated using the patient’s age and normalized percentile equations published by Bonafide et al 2013. To exclude readings likely due to artifact, thresholds for heart rate, respiratory rate, and oximetry values were also applied (see Additional Methods).

2.4 | Statistical analysis

Patients were randomly assigned into a training (80%) or testing (20%) set (balanced validation) and compared with ensure balance by demographic criteria and clinical criteria. A separate randomized 10-fold cross validation was conducted to further validate findings.
2.5 Machine learning

Supervised machine learning techniques use input and output data to find patterns and make predictions. A cascaded ANN was used to predict a respiratory score ranging from 1 to 15, based on training inputs. Neural networks depend on linking "neurons" or multiple learning units to detect patterns in data. In comparison with the conventional feed-forward neural network, the cascaded network structure is more advanced. It augments a set of cascaded paths to direct the nodes in the preceding and current layers to be the input into the next layer (Figure S3). The cascaded ANN included eight hidden layers with 3 to 50 neurons in each layer. Machine learning regression models based on Normal and Poisson distribution were used for comparative purposes. The accuracy of machine learning models was assessed by comparison of the median absolute error (MAE) for each of the testing sets. Matlab (version R2017a) was the program used to create the machine learning algorithm.

3 RESULTS

The health record query returned 186 eligible patients admitted to the PICU for status asthmaticus without a potential confounding diagnosis as described in our exclusion criteria. Of the 18 621 patients were excluded due to lack of stored vital sign data during the admission and an additional 37 patients were excluded due to lack of PAS scores that aligned with available extracted vital sign data. The remaining N = 128 patients were included in our randomization and subsequent machine learning analysis (Figure 1).

A total of 50.8% of the inclusion cohort were male, with 34.4% Hispanic/Latino ethnicity, and 19.5% identifying as African American/Black. The median age was 7.9 years old and the median length of hospital stay was 79 hours (Table 1).

In our training data, we had 4943 original PAS scores or 12.5% of the recorded PAS for the balanced testing set. The results of the balanced group (n = 102 patients and 34 000 to 38 000 data points) was reduced slightly to 36 321 data points after application of the artifact thresholds for age and respiratory rate. For the 10-fold cross-validation, the number of data points included in each fold of training varied from approximately 34 000 to 38 000.

On the basis of the comparison of MAE for the balanced testing set for each of the machine learning models, the cascaded ANN with eight hidden layers trained with the balanced group yielded the smallest MAE of 1.21. The MAEs across the balanced group Poisson and Normal models were 1.24 and 1.25, respectively. The Poisson and Normal models each yielded slightly higher MAEs for the extreme values of the PAS scores (Figure 2). The most accurate predictions occurred in the mid-range values of 6-9, where the most training data existed (Figure 2).

In our asthma clinical care guidelines, 2-point discrimination on the PAS scale is a clinically relevant range for guiding care. Thus, we also evaluated pARS predictability in the ±2 point range. Specifically, 80% of the pARS scores produced by the ANN algorithm are within ±2.10 of the recorded PAS for the balanced testing set. The results from the 10-fold cross validation are similar (Figure S5).

The pARS values are also aligned well with PAS when mapped over time across the course of individual patient encounters (Figure 3).

---

**TABLE 1** Cohort description and demographics, distribution of demographics within balanced testing and training sets, reported race, and ethnicity

<table>
<thead>
<tr>
<th>Race</th>
<th>Full cohort (n = 128)</th>
<th>Training (n = 102)</th>
<th>Test (n = 26)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male)</td>
<td>65 (50.8%)</td>
<td>50 (49.0%)</td>
<td>15 (57.7%)</td>
</tr>
<tr>
<td>Age, y</td>
<td>7.9 (4.4-11.4)</td>
<td>7.6 (4.0-11.4)</td>
<td>9.0 (5.4-11.4)</td>
</tr>
<tr>
<td>Hispanic or Latino ethnicity</td>
<td>44 (34.4%)</td>
<td>34 (33.3%)</td>
<td>10 (38.5%)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>66 (51.6%)</td>
<td>51 (50.0%)</td>
<td>15 (57.7%)</td>
</tr>
<tr>
<td>Black/African American</td>
<td>25 (19.5%)</td>
<td>19 (18.6%)</td>
<td>6 (23.1%)</td>
</tr>
<tr>
<td>Other*</td>
<td>37 (28.9%)</td>
<td>32 (31.4%)</td>
<td>5 (19.2%)</td>
</tr>
<tr>
<td>Length of stay, h</td>
<td>79 (42-117)</td>
<td>71 (39-112)</td>
<td>85 (50-138)</td>
</tr>
<tr>
<td>Max PAS score</td>
<td>10 (9-12)</td>
<td>10 (9-12)</td>
<td>11 (10-12)</td>
</tr>
<tr>
<td>PAS score, median (range)</td>
<td>7 (5-15)</td>
<td>7 (5-15)</td>
<td>8 (5-13)</td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; PAS, pediatric asthma score. *Other race includes: American Indian/Alaska Native (n = 1), Asian (n = 1), Native Hawaiian/Other Pacific (n = 1), unknown/not reported (n = 3), other (n = 23) and more than one race (n = 9).
4 | DISCUSSION

We successfully created pARS, a novel, pediatric-automated asthma severity score, using physiologic data and machine learning. Using an ANN trained with three vital sign parameters, the pARS was well within 2 points of recorded PAS scores based on the analysis of MAE. This level of accuracy makes the automated score noninferior to the manual PAS used at our institution.

The pARS was most accurate in the mid-range of asthma severity between 6 and 9. This severity range is when critical decisions about patient care (transfer to ICU vs floor) are made, increasing the potential positive impact of an objective decision support tool that uses pARS.

This study is foundational in its use of the secondary electronic medical data merged with the passive vital sign monitoring information collected in the course of patient care to create an automated pediatric respiratory score. The strengths of this machine score are its objectivity, and that it uses data collected automatically from monitors already in use. It also incorporates age-based parameters. It is unique in its ability to continuously monitor acute changes in asthma status with a computed score.

Because it is objective, automated, and can be continuously generated, this score has the potential to help standardize acute pediatric asthma care in the PICU. Studies of PICU management of severe asthma across pediatric hospitals and even within a single institution have revealed marked variability in practice. While there are global treatment guidelines for asthma in the primary care and emergency department setting, acute care of life-threatening pediatric asthma, particularly in severe exacerbations treated in the PICU, is not standardized. Respiratory scoring systems for pediatric asthma have streamlined and improved inpatient care, but there is an inconsistent use of any such scoring system in the PICU. Bartlett et al implemented and studied a successful clinical care guideline in the PICU, which included a bronchodilator weaning pathway based on a subjective staff generated respiratory score called the modified pulmonary index score and showed a decreased length of stay in the hospital overall. An automated, quantitative score to replace these respiratory scores is appealing to critical care providers who are concerned about evidence-based clinical decision making, interrater reliability, and staff efficiency.

Currently, no automated asthma severity score has been published that is sensitive to changes in acute respiratory status. Existing work in pediatric asthma prediction has focused on predicting the occurrence of asthma exacerbations and asthma control deterioration. Other machine learning and predictive analytic work in pediatrics have targeted sentinel events such as sepsis, respiratory failure requiring intubation, and cardiac arrest as outcomes. The frequency of these events is low and thus requires a large patient population and significant monitoring time to assess validity. The pARS score can be validated over a shorter time frame and in smaller populations because it assesses and learns the continuum of disease severity and can be associated with more common outcomes, for example,
decreasing length of stay in PICU rather than leveraging outcomes such as respiratory arrest and death that are discrete rare events.

Limitations of this study include incomplete data from the clinical record and vital sign database, which decreased our inclusion cohort. While data fidelity issues are innate to using electronic clinical record data for secondary research investigation, this highlights the importance of considering issues of data input and integrity as we build our health data systems, particularly as it becomes imperative to merge and map multiple data sources. Because of the practice patterns in our PICU, the PAS score is not recorded as frequently as on the inpatient floors, which limited data analysis, further emphasizing the need for an automated score. In addition, our algorithm training was restricted to the specific PAS score used at our center and therefore may not generalize to systems that use other versions of manual clinical scoring. We were able to build this score using a heterogenous pediatric patient population at altitude (~5280 ft) seen at an urban quaternary care center. All patients included were on oxygen while in the PICU as a part of the standard of care. This introduces variability into the pulse oximetry data incorporated into the algorithm. Future prospective, multicenter study is needed to further elucidate the impact of oxygen therapy on vital sign parameters.

Our study algorithm was also not designed to include other clinical features or risk factors that can influence asthma severity. For example, acute bronchodilator use, the effect of noninvasive ventilation, as well as markers of chronic asthma severity including controller medication use and adherence, atopy, prior hospitalization history, recent exacerbations, symptom surveys, baseline FEV$_1$%, and environmental exposures could all impact pARS predictive ability. Prospective collection of some of these parameters in future studies may strengthen our model and allow further personal risk stratification and treatment algorithms. Work by Luo et al$^{23}$ in 2015 was able to merge 2 years of pediatric asthma personal symptoms data collected via a mobile application with patient attributes and environmental variables to successfully predict a child’s asthma control deterioration 1 week ahead.

Our study analyzed vital sign parameters measured at 1-minute intervals. Because of the limited computing power and data storage and access limitations, we were not able to analyze the dense data contained in these vital sign waveforms. The ability to apply deep or unsupervised machine learning to high data density vital sign waveforms has enormous predictive potential. For example, trends in heart rate variability have been used to predict neonatal sepsis and arterial waveforms analyzed to predict hypovolemic shock.$^{26-28}$ Advances in mobile and hospital-based monitoring and cloud-based analytics are innovations that will help facilitate research and clinical application of high-resolution physiologic data, conquering current challenges of poor fidelity data due to artifact and variable collection, and the size of the data files.

Any decision support tool is only useful if providers trust and utilize it. Implementation and adoption of this tool in the PICU will require significant changes to workflow with patients and the use of the EHR. Significant quality improvement and culture change work will be needed to adopt this scoring system as a clinical decision-making tool.

Given these limitations and context, the next steps are to work prospectively to validate pARS as a research instrument and associate it with meaningful clinical outcomes first in our institution, and then more broadly. Technologic innovation addressing issues including integration of data sources and the EHR, computing power, and the speed to run machine learning methods continuously will be required to translate this score into a clinician decision support tool. With prospective validation, real-time implementation and workflow adoption, we believe a score like pARS can drive higher quality care, improve patient flow, decrease the length of stay and medication burden (facilitate timely weaning of continuous medications).

5 | CONCLUSION

This study shows that the creation of a pARS leveraging machine learning techniques such as ANNs to analyze simple vital sign parameters and limited clinical data is feasible. The potential impact of such a score to improve and standardize the PICU management of acute asthma exacerbation is significant. Our study revealed multiple barriers to integration of disparate clinical data sources and was also weakened by incomplete EHR and monitor data, both common challenges in studies using secondary retrospective queries on data produced by routine clinical practice. Further prospective validation of our algorithm is imperative to improve data integrity, refine and expand contributing features, and assess the impact of pARS on clinical outcomes including length of stay, and medication burden and operational impacts such as staff efficiency.

ACKNOWLEDGMENTS

We would like to acknowledge the contributions of Alix Mizoue, Sarah Deakyne, Jennifer Sadlowski, and Deborah Batson for their assistance with data extraction. Kyle Robison in assisting with COMIRB project submission; Dr Paul Stillwell for his clinical editing skills; as well as Dr Angela Czaja for contributing her expertise and insights specific to the pediatric ICU.

AUTHOR CONTRIBUTIONS

NB: Patents: (1) Apparatus and methods for smartphone-based Sp02 measurement; (2) Respiratory Rate Measurement System; (3) TYTH: Typing On Your Teeth: Tongue-Teeth Localization for Human-Computer Interface.

SJS has consulted for Aerocrine, AstraZeneca, Boehringer Ingelheim, Daiichi Sankyo, GlaxoSmithKline, Genentech, Merck, Novartis, Propeller Health, Regeneron, Roche, Sanofi, and Teva.

TV was Cofounder and Board Member of Now Vitals, Inc, Cofounder and President Earable Inc. Patents: (1) Method, apparatus, and system for capacitive touch communication; (2) Thermal comfort building monitoring with wearables: method, apparatus, and system; (3) Resource Access Permission Recommendation Method for Mobile Applications; (4) Method, apparatus, and system for two factor authentication in single step with wearable devices; (5) Method, apparatus, and system for noncontact breathing volume monitoring.

**RR Deterding:** Cofounder and Board Member—Triple Endoscopy, Inc. (Patent Nasal Endoscopy Scope), Advisory Board Pediatric Interstitial Lung Disease for Boehringer Ingelheim, Cofounder and President Now Vitals, Inc, Cofounder and Board Member Earable Inc. Patents: (1) Computing Systems for Determining Vital Information; (2) Personalized Health Care Wearable Sensor System; (3) Respiratory rate measurement system; (4) Health Sticker: A modular adhesive platform monitoring vital signs; (5) Breathing Gripper: A miniature breath monitoring device.

**FUNDING INFORMATION**

AIM received funds by NIH and NHLBI, Abman T-32 Fellowship Training Grant no. #2T32HL007670-26. SJS received support from the Colorado Department of Public Health and Environment, and was also supported by the NHLBI AsthmaNet U10 HL098075 and the National Heart, Lung, and Blood Institute, 1UG1HL139123, Precision Interventions for Severe and/or Exacerbation Prone Asthma (PrecISE), TV received support from the National Science Foundation, Department of Homeland Security, Colorado Office of Economic Development and International Trade, University of Colorado, Google Inc, Microsoft Inc, and Qualcomm Inc. RR D received support from the Colorado Advanced Industries Accelerator (AIA).

**REFERENCES**


SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Messinger AI, Bui N, Wagner BD, Szefler SJ, Vu T, Deterding RR. Novel pediatric-automated respiratory score using physiologic data and machine learning in asthma. *Pediatric Pulmonology*. 2019;1-7. https://doi.org/10.1002/ppul.24342