



## CLINICAL RESEARCH ARTICLE

# Vital sign metrics of VLBW infants in three NICUs: implications for predictive algorithms

Amanda M. Zimmet<sup>1</sup>, Brynne A. Sullivan<sup>2</sup>, Karen D. Fairchild<sup>2</sup>, J. Randall Moorman<sup>1</sup>, Joseph R. Isler<sup>3</sup>, Aaron W. Wallman-Stokes<sup>3</sup>, Rakesh Sahni<sup>3</sup>, Zachary A. Vesoulis<sup>4</sup>, Sarah J. Ratcliffe<sup>5</sup> and Douglas E. Lake<sup>1</sup>

**BACKGROUND:** Continuous heart rate (HR) and oxygenation (SpO<sub>2</sub>) metrics can be useful for predicting adverse events in very low birth weight (VLBW) infants. To optimize the utility of these tools, inter-site variability must be taken into account.

**METHODS:** For VLBW infants at three neonatal intensive care units (NICUs), we analyzed the mean, standard deviation, skewness, kurtosis, and cross-correlation of electrocardiogram HR, pulse oximeter pulse rate, and SpO<sub>2</sub>. The number and durations of bradycardia and desaturation events were also measured. Twenty-two metrics were calculated hourly, and mean daily values were compared between sites.

**RESULTS:** We analyzed data from 1168 VLBW infants from birth through day 42 (35,238 infant-days). HR and SpO<sub>2</sub> metrics were similar at the three NICUs, with mean HR rising by ~10 beats/min over the first 2 weeks and mean SpO<sub>2</sub> remaining stable ~94% over time. The number of bradycardia events was higher at one site, and the duration of desaturations was longer at another site.

**CONCLUSIONS:** Mean HR and SpO<sub>2</sub> were generally similar among VLBW infants at three NICUs from birth through 6 weeks of age, but bradycardia and desaturation events differed in the first 2 weeks after birth. This highlights the importance of developing predictive analytics tools at multiple sites.

*Pediatric Research* (2021) 90:125–130; <https://doi.org/10.1038/s41390-021-01428-3>

**IMPACT:**

- HR and SpO<sub>2</sub> analytics can be useful for predicting adverse events in VLBW infants in the NICU, but inter-site differences must be taken into account in developing predictive algorithms.
- Although mean HR and SpO<sub>2</sub> patterns were similar in VLBW infants at three NICUs, inter-site differences in the number of bradycardia events and duration of desaturation events were found.
- Inter-site differences in bradycardia and desaturation events among VLBW infants should be considered in the development of predictive algorithms.

**INTRODUCTION**

Vital signs that reflect the cardiovascular and respiratory systems are continuously displayed on bedside monitors in the neonatal intensive care unit (NICU), and aberrations may signal a variety of pathologic processes.<sup>1</sup> Subtle changes can occur before overt clinical signs of illness, prompting the development of early warning systems that alert clinicians to changes in patient status requiring attention.<sup>2–4</sup> One example is the finding of abnormal heart rate (HR) characteristics of decreased variability and transient repetitive decelerations that sometimes precede the clinical presentation of sepsis, necrotizing enterocolitis (NEC), or other infections in very low birth weight (VLBW) preterm infants.<sup>5–9</sup>

In a nine-NICU randomized clinical trial of 3003 VLBW infants, display of an HR characteristic index, the fold increase in risk of sepsis being diagnosed in the next day, was associated with a 22% relative decrease in mortality rate.<sup>10</sup> Another example of a change in vital signs in preterm infants is the simultaneous fall in HR and

oxygen saturation (SpO<sub>2</sub>) during neonatal apnea, the familiar bradycardia-desaturation spell.<sup>11–13</sup> A measure of this, the maximum cross-correlation of HR and SpO<sub>2</sub>, increased prior to diagnosis of sepsis or NEC in a study of 1065 VLBW infants in two NICUs.<sup>14</sup>

HR and SpO<sub>2</sub> are affected not only by illness and stress but also by maturation and by clinical care practices such as the mode of respiratory support.<sup>15–17</sup> Different bedside monitor hardware and sensors may also contribute to differences in vital sign measurements across units. Here, we examined the ranges of values of canonical vital signs for >1000 VLBW infants at three large tertiary care NICUs during the first 6 weeks of hospitalization. We also compared the number of bradycardia and desaturation events and the cross-correlation of HR and SpO<sub>2</sub>. As a step toward developing mathematical predictive algorithms that are generalizable across NICUs, we sought to determine the expected ranges of these parameters over time and how they varied among infants at the three sites.

<sup>1</sup>Department of Medicine, University of Virginia, Charlottesville, VA, USA; <sup>2</sup>Department of Pediatrics, University of Virginia, Charlottesville, VA, USA; <sup>3</sup>Department of Pediatrics, Columbia University, New York, NY, USA; <sup>4</sup>Department of Pediatrics, Washington University St. Louis, St. Louis, MO, USA and <sup>5</sup>Department of Public Health Science, University of Virginia, Charlottesville, VA, USA

Correspondence: Karen D. Fairchild (kdf2n@virginia.edu)

Received: 31 July 2020 Revised: 3 February 2021 Accepted: 4 February 2021

Published online: 25 March 2021

**METHODS**

We analyzed vital sign data from VLBW infants ( $\leq 1500$  g birth weight) admitted from 2012 to 2018 at three-level IV NICUs (University of Virginia (UVA): University of Virginia Children’s Hospital, Columbia University (CU): NewYork-Presbyterian Morgan Stanley Children’s Hospital, and Washington University in St. Louis (WUSTL): St. Louis Children’s Hospital). Institutional Review Boards at each site approved the study with a waiver of consent. We excluded infants with congenital or chromosomal anomalies that could impact oxygenation, those transitioned to comfort care only, and those with fewer than 7 days of HR and SpO<sub>2</sub> data within the first 28 days after birth.

The three participating centers routinely collect and store NICU bedside monitor vital sign data using the BedMaster system (Excel Medical, Jupiter, FL). In addition, UVA collects data using the Cardiopulmonary Corporation system (Milford, Connecticut). During the period of study, UVA and CU NICUs used GE bedside monitors (GE Healthcare, Waukesha, WI) with Masimo pulse oximeters (Masimo Corporation, Irvine, CA), and data were recorded at 0.5 Hz. The WUSTL NICU used Philips monitors (Philips Corporation, Andover, MA) with Nellcor Oximax pulse oximeters (Medtronic, Minneapolis, MN), and data were recorded at 1 Hz but down-sampled to 0.5 Hz to match the other sites. All pulse oximeters had an 8 s averaging time. During the study period, UVA and WUSTL clinicians had a default goal SpO<sub>2</sub> range 88–95%, increasing slightly as infants approached term-corrected gestational age. CU used a goal range of 85–93% until August 2013 and then switched to 90–95%. Bradycardia alarms were set at 90 beats per minute (b.p.m.) at UVA and 100 b.p.m. at the other two sites.

**HR, PR, and SpO<sub>2</sub> metrics**

We analyzed continuously measured electrocardiogram-derived HR, pulse oximeter-derived pulse rate (PR), and SpO<sub>2</sub>. Daily mean, standard deviation, skewness, and kurtosis of HR, PR, and SpO<sub>2</sub> were computed for each infant over the first 6 weeks after birth. To control for artifact, all values of zero were removed and, for measurements other than mean, values >99th percentile were censored to the 99th percentile value.

Bradycardia and desaturation events were quantified using thresholds and definitions we have previously published.<sup>18,19</sup> Bradycardia was defined as HR <100 b.p.m. for at least 4 s and desaturation as SpO<sub>2</sub> < 80% for at least 10 s. Events were joined if they occurred within 4 or 10 s of each other for bradycardia and desaturation, respectively. We report the mean number and duration of events per day as well as the percentage of time spent in bradycardia or desaturation. We calculated the cross-correlation of HR or PR and SpO<sub>2</sub>%. We used our own code written in Matlab for the analyses. Data were smoothed using a sliding window of 7 days as we have done in prior work.<sup>20,21</sup>

**Statistics**

We assessed for site effects on each metric using daily means from the day of birth through day 42 by *n*-way analysis of variance. Pairwise comparisons between sites used a Bonferroni correction to account for multiple comparisons, with significance set at  $p < 0.05/42/3$  (42 days of comparisons, three pairwise comparisons). Figures show estimated marginal means corrected for birth weight, gestational age, and sex differences between sites. Estimated population marginal means control for the influence of the covariates (gestational age, birthweight, and sex) on the outcome variable of interest (HR, SpO<sub>2</sub>%, etc.).<sup>22</sup> They adjust for any biases from imbalances in the covariates. The estimated mean for the variable of interest is based on the equal-weighting method, resulting in adjusted means that are equally balanced across all values of all covariates. To calculate the estimated marginal means, we used the multcompare function in Matlab using a linear repeated-measures model of the data from the anovan function. The statistical impact of the site on a particular

**Table 1.** Demographic and clinical variables.

	UVA	CU	WUSTL
<i>N</i>	477	444	247
Median birth weight (g) (IQR)	1030 (798–1280)	1016 (763–1285)	935 (718–1150)
Median gestational age (weeks) (IQR)	28 (25–30)	28 (26–30)	27 (25–28)
% Female	50%	56%	51%
Median days of data per infant (IQR)	39 (26–42)	33 (20–41)	19 (11–39)
Total days of data	16,137	13,305	5796
Median length of stay (days) (IQR)	61 (39–96)	63 (41–95)	81 (56–117)
Mortality	3%	4%	15%

metric was measured using log<sub>10</sub>(*p* value), that is, by reporting the number of leading zeros for the *p* value.

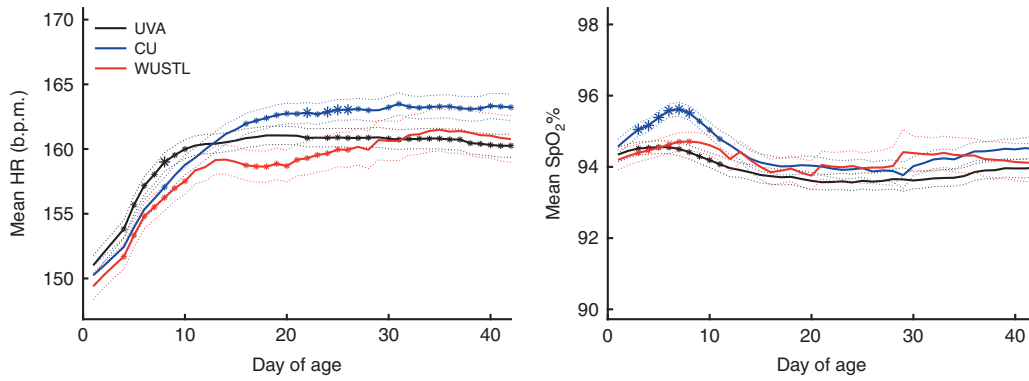
**RESULTS**

During the period of study, 3209 VLBW infants were admitted to the three NICUs with vital sign data recording available, 1168 of whom had no exclusions and had at least 7 days of stored vital sign data available for analysis in the first 4 weeks after birth. Demographics of the infants in the three site cohorts are shown in Table 1. We analyzed 35,238 infant-days of data (96 infant-years). The distribution of data availability by postmenstrual age (PMA) was the same for UVA and CU, but WUSTL had lower coverage after 28 weeks PMA (Supplementary Figure S1).

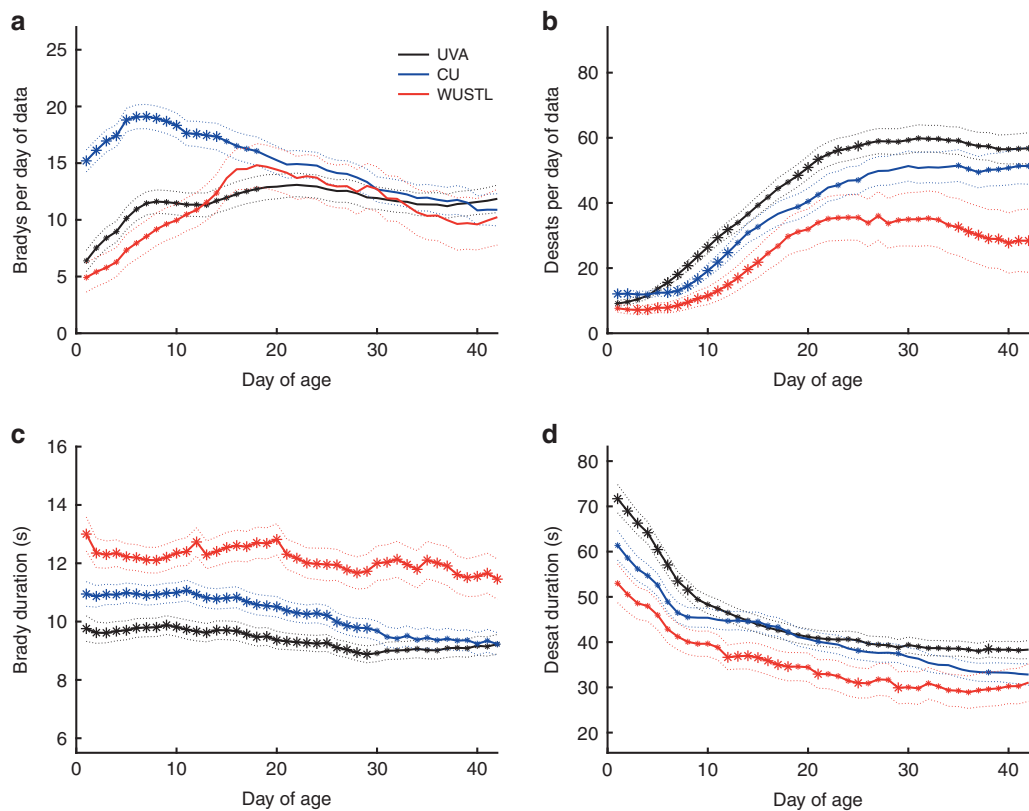
As shown in Fig. 1, the mean HR and SpO<sub>2</sub>% were similar at the three sites over the 6 weeks of study. The mean HR rose from ~150 b.p.m. in the first week to ~160 b.p.m. and changed little thereafter. After 2 weeks of age, there was a small (~4 b.p.m.) difference in infants’ daily mean HR between sites. The daily mean SpO<sub>2</sub> was slightly different (~1%) between sites in the first 2 weeks after birth.

Figure 2 shows the number (top panels) and durations (bottom panels) of bradycardia events (left) and desaturation events (right). The differences were as large as twofold; infants at CU had up to twice as many bradycardia events per day, and infants at WUSTL had about half as many desaturation events, with the magnitude of the differences varying over time. By 3 weeks after birth, the difference in daily numbers of bradycardias between sites was no longer evident, while the difference in daily numbers of desaturations between sites increased from birth to 6 weeks. The smaller differences in event durations remained similar throughout. The percentages of time spent in bradycardia or desaturation are shown in Supplementary Figure S2. The number of bradycardia and desaturation events are shown split by birthweight in Supplementary Figure S3.

Although the absolute differences in some of the HR and SpO<sub>2</sub> metrics between sites were very small, the large number of data points analyzed gave some of these differences high statistical significance. This is depicted in Fig. 3 as a heat map of the number of leading zeros in the *p* value for inter-site differences in each metric for each day from birth through day 42 (with correction for multiple comparisons, thus statistical significance set at  $p < 0.05/42$  or approximately  $p < 0.001$ ). Metrics are ordered from those with the most to the least inter-site differences. Notably, skewness of PR measured from the pulse oximeter had more significant inter-site differences (appearing near the top of the list of metrics) compared to skewness of HR measured from the electrocardiogram (appearing near the bottom of the list). Individual trends for all Fig. 3 metrics not shown in Figs. 1 and 2 are shown in Supplementary Figures S4–8. Supplementary Figure S9 provides a probability density plot for all vital sign metrics in Fig. 3.



**Fig. 1** HR and SpO<sub>2</sub> trends of VLBWs at three NICUs. Daily mean HR (left) and SpO<sub>2</sub> (right) are shown for VLBW infants at the three NICUs through the first 6 weeks from birth. Thin dotted lines indicate the 95% confidence interval. Asterisks indicate a statistically significant difference on that day compared to one other site (small asterisk) or both other sites (large asterisk). Y-axis limits are the 10th and 90th percentiles.



**Fig. 2** Bradycardia and desaturation events by site. Mean number of bradycardia (a) and desaturation (b) events per day of data are shown for the first 6 weeks from birth for VLBW infants at UVA, CU, and WUSTL. Mean event duration in seconds is shown in the bottom panels (c, d). Dotted lines indicate the 95% confidence interval. Asterisks indicate a statistically significant difference on that day compared to one other site (small asterisk) or both other sites (large asterisk). Y-axis limits are the 10th and 90th percentiles.

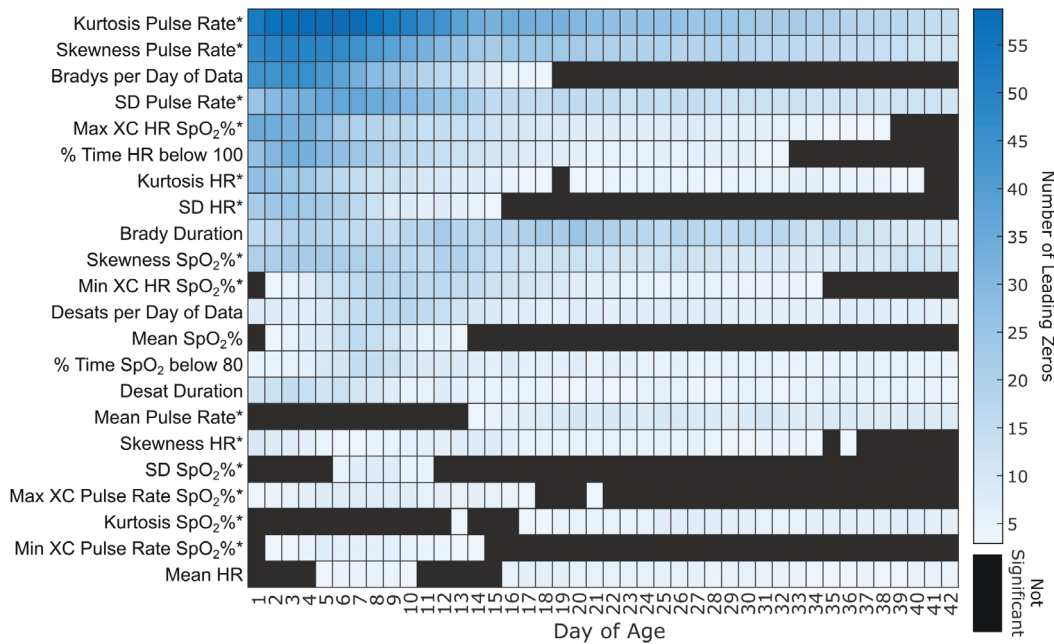
Using the average value for each infant for all HR, PR, and SpO<sub>2</sub>% metrics across each infant's whole stay, we ran a rank-sum test to look for a difference between sexes. Upon correcting for birthweight, gestational age, and institution, only the mean, skewness, and kurtosis of SpO<sub>2</sub>% were significantly different ( $p < 0.05$ ) between the sexes (Supplementary Figure S10), but the differences were small (<1% difference in mean SpO<sub>2</sub>%).

**DISCUSSION**

Abnormal values, trends, and patterns of continuously monitored vital signs in NICU patients can predict imminent or longer-term

adverse events and outcomes. Assessment of potential inter-site differences in infants' vital sign patterns is needed in order to optimize predictive algorithms. We, therefore, performed a three-center comparison of the most frequently monitored vital signs in VLBW infants, HR (HR measured by ECG and PR measured by pulse oximeter) and SpO<sub>2</sub>, in the first 6 weeks after birth. We found inter-center variability that may reflect differences in patient populations, equipment, or care practices.

With regard to HR and SpO<sub>2</sub>, Fig. 1 shows that the overall mean HR increased from ~150 b.p.m. in the first week after birth to ~160 b.p.m. from weeks 2–6, while the mean SpO<sub>2</sub> of ~94% was consistent over the time period studied. The change in HR over



**Fig. 3 Magnitude of statistical significance of site differences in HR, PR, and SpO<sub>2</sub> metrics.** For each metric shown on the left y-axis, the number of zeros preceding the *p* value for inter-site differences each day shown on the x-axis is depicted as a heat map. Black boxes indicate no significant difference between sites (adjusting for 42 comparisons,  $p < 0.05/42$  or approximately  $p < 0.001$ ). Progressively darker shades of blue indicate more leading zeros in the *p* value for inter-site differences. \*Computed using 10 min averages. SD standard deviation, XC cross-correlation, HR ECG heart rate.

time that we show in this VLBW cohort is similar to that previously published for term infants, but with an offset of ~20 b.p.m. (preterm infants have higher HR than term infants).<sup>23</sup> The HR values in Fig. 1 are also similar to those previously published in a single site report on preterm infants at UVA.<sup>24</sup> There were inter-site differences of ~4 b.p.m. in HR and 1% in SpO<sub>2</sub>, which, due to the large volume of data, were statistically significant, if not clinically meaningful. Whether these small differences would impact a mathematical model to predict outcomes would be model-specific; this highlights the importance of developing and testing models at multiple sites.

We also found inter-site differences in bradycardia and desaturation events. In Fig. 2, we note that infants at CU had more bradycardia events during the first 2 weeks after birth. Our definition of bradycardia of <100 b.p.m. was the same as the alarm threshold at CU and WUSTL (and only 10 b.p.m. higher than the alarm threshold at UVA) and thus it is unlikely that the difference in the number of bradycardia events is due to center-specific alarm management. A possible explanation for more bradycardia events at CU is less use of mechanical ventilation and more use of nasal continuous positive airway pressure<sup>25</sup> leading to more apnea-associated HR decelerations.<sup>26</sup> This is also supported by higher cross-correlation of HR and SpO<sub>2</sub> (Supplementary Figure S8). With regards to desaturations, we found lower rates and durations for infants at WUSTL compared to the other two sites. The reason for this difference is unknown, but may relate to the monitor alarm tones. UVA and CU use monitors with a high alert tone for bradycardia events and a softer alert tone for desaturation events, whereas WUSTL monitors give the same alert tone for both desaturations and bradycardias. Another consideration is that different monitors and sensors have different hardware and algorithms, which could impact vital sign values. We are not implying the bradycardias and desaturations are benign; we are highlighting that differences in clinical care and patient populations between NICUs can impact bradycardias and desaturations. Therefore,

cardiorespiratory predictive algorithms should be externally validated.

The small but statistically significant difference in cross-correlation of HR and SpO<sub>2</sub> between sites, especially in the first week after birth (Supplementary Figure S7) may be an important finding since we identified its association with apnea and exaggerated periodic breathing.<sup>13</sup> Moreover, the cross-correlation of HR and SpO<sub>2</sub> was a significant predictor in a model targeting imminent septicemia or NEC.<sup>14</sup> In that study of >1000 VLBW infants, we also found that infants at CU had a slightly higher baseline cross-correlation of HR and SpO<sub>2</sub> than infants at UVA. The mechanism is unknown, but may relate to less mechanical ventilation at CU and thus more apnea, with a concurrent decline in HR and SpO<sub>2</sub>.

The strength of this analysis is the large number of VLBW infants and days of data analyzed at three NICUs with diverse patient populations and clinical practices. We acknowledge there are a number of limitations as well. We do not have individual patient data on daily respiratory support in the infants included in these analyses to validate the assumption that different approaches to mechanical ventilation at the three units impact desaturation and bradycardia events due to apnea. We are able to report more generally, however, that days on mechanical ventilation for VLBW infants is quite different at CU compared to WUSTL and UVA (mean 10, 35, and 33 days, respectively, in 2017–2018). Another limitation is that we do not have dates and doses of caffeine, although practices for caffeine administration are similar at the three NICUs. Also, the patient demographics and outcomes are different at WUSTL compared to the other two sites in that infants were, on average, ~1 week lower gestational age and had higher morbidities and mortality. This likely reflects the sociodemographic variables that contribute to well-documented higher infant mortality in St. Louis compared to the other two sites.<sup>27–29</sup> In the future, we will address the impact of mechanical ventilation and oxygen support on cardiorespiratory events and outcomes of extremely preterm infants in the

Pre-Vent multi-NICU collaboration, in which there are granular data on daily respiratory support, medications, and clinical outcomes linked to bedside monitor vital sign data on over 700 infants <29 weeks gestation.<sup>30</sup>

The differences we see here highlight the importance of multicenter studies, especially when developing predictive analytics. Variations in demographics, clinical practices, and monitors or sensors all have an impact on continuous vital sign data. More than 40 years ago, Ransohoff and Feinstein<sup>31</sup> analyzed why diagnostic tests fail. They advanced the concept, later called *spectrum bias*, that a test is limited if developed on diseased patients who did not represent the spectrum of pathology or clinical features, or if tested on control patients that had a different spectrum of comorbidity. A vivid example of failed external validation of not one but dozens of predictive models is the recent Physionet Sepsis Challenge.<sup>32</sup> No model that had a good performance on the two-hospital training data set did at all well on a test set from a third hospital. We note, however, a prominent example of a successful NICU predictive model generated at a single center, the HR characteristics index developed at UVA, which performed well in external validation at a second NICU and was then shown to reduce mortality in a nine-NICU study.<sup>10</sup> Moreover, in more recent work, an HR and SpO<sub>2</sub> model for predicting sepsis performed well at both UVA and CU in spite of differences in vital sign trends.<sup>14</sup>

## CONCLUSION

In a three-NICU study of 1168 VLBW infants from birth through 6 weeks of age, we found that mean HR and SpO<sub>2</sub> were generally similar, but bradycardia and desaturation events differed in the first 2 weeks after birth. The differences we found in bradycardia and desaturation events between sites may inspire mechanistic studies into the impact of variations in respiratory support or other clinical practices on measures of cardiorespiratory instability, which may impact clinical outcomes. Since this work is presented here in the context of developing tools for predictive analytics monitoring, our findings highlight the importance of developing and validating vital sign-based analytics at multiple sites.

## DATA AND CODE

Summary data and code are available on UVA Dataverse at <https://doi.org/10.18130/V3/7UAPHU>.

## ACKNOWLEDGEMENTS

This work was supported by HHS|NIH|Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) R01 HD072071 and K23 HD097254-01. HHS|NIH|National Institute of Neurological Disorders and Stroke (NINDS) K23 NS111086.

## AUTHOR CONTRIBUTIONS

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data, and drafting the article or revising it critically for important intellectual content. All authors gave final approval of the version to be published.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41390-021-01428-3>.

**Competing interests:** J.R.M. and D.E.L. have equity shares in Medical Predictive Science Corporation, Charlottesville, VA. J.R.M. is an officer and owns equity in Advanced Medical Predictive Devices, Diagnostics, and Displays. S.J.R. has received consulting fees from Airway Therapeutics. The other authors declare no competing interests.

**Patient consent:** Institutional Review Boards at each site approved the study with waiver of consent.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

- Kumar, N., Akangire, G., Sullivan, B., Fairchild, K. & Sampath, V. Continuous vital sign analysis for predicting and preventing neonatal diseases in the twenty-first century: big data to the forefront. *Pediatr. Res.* <https://doi.org/10.1038/s41390-019-0527-0> (2019).
- Joshi, R. et al. Predicting neonatal sepsis using features of heart rate variability, respiratory characteristics, and ECG-derived estimates of infant motion. *IEEE J. Biomed. Heal. Inform.* **24**, 681–692 (2020).
- Dewhurst, C. J., Cooke, R. W. I. & Turner, M. A. Clinician observation of physiological trend monitoring to identify late-onset sepsis in preterm infants. *Acta Paediatr. Int. J. Paediatr.* **97**, 1187–1191 (2008).
- Mithal, L. B. et al. Vital signs analysis algorithm detects inflammatory response in premature infants with late onset sepsis and necrotizing enterocolitis. *Early Hum. Dev.* **117**, 83–89 (2018).
- Sullivan, B. A. & Fairchild, K. D. Predictive monitoring for sepsis and necrotizing enterocolitis to prevent shock. *Semin. Fetal Neonatal Med.* **20**, 255–261 (2015).
- Griffin, M. P., Moorman, J. R., Todd, J., Workman-Daniels, K. & Goldmann, D. Toward the early diagnosis of neonatal sepsis and sepsis-like illness using novel heart rate analysis. *Pediatrics* **107**, 97–104 (2001).
- Griffin, M. P. et al. Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-like illness. *Pediatr. Res.* **53**, 920–926 (2003).
- Griffin, M. P. et al. Heart rate characteristics: novel physiologic markers to predict neonatal infection and death. *Pediatrics* **116**, 1070–1074 (2005).
- Weitkamp, J. H. et al. Meningitis, urinary tract, and bloodstream infections in very low birth weight infants enrolled in a heart rate characteristics monitoring trial. *Pediatr. Res.* **87**, 1226–1230 (2020).
- Moorman, J. R. et al. Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial. *J. Pediatr.* **159**, 900–907 (2011).
- Lee, H. et al. A new algorithm for detecting central apnea in neonates. *Physiol. Meas.* **33**, 1–17 (2012).
- Clark, M. T. et al. Breath-by-breath analysis of cardiorespiratory interaction for quantifying developmental maturity in premature infants. *J. Appl. Physiol.* **112**, 859–867 (2012).
- Fairchild, K. D. & Lake, D. E. Cross-correlation of heart rate and oxygen saturation in very low birthweight infants: association with apnea and adverse events. *Am. J. Perinatol.* **35**, 463–469 (2018).
- Fairchild, K. D. et al. Vital signs and their cross-correlation in sepsis and NEC: a study of 1,065 very-low-birth-weight infants in two NICUs. *Pediatr. Res.* **81**, 315–321 (2017).
- Martin, R. J., DiFiore, J. M., Macfarlane, P. M. & Wilson, C. G. Physiologic basis for intermittent hypoxic episodes in preterm infants. *Adv. Exp. Med. Biol.* **758**, 351–358 (2012).
- Bancalari, E. & Claire, N. Respiratory instability and hypoxemia episodes in preterm infants. *Am. J. Perinatol.* **35**, 534–536 (2018).
- McDonald, F. B., Dempsey, E. M. & O'Halloran, K. D. The impact of preterm adversity on cardiorespiratory function. *Exp. Physiol.* **105**, 17–43 (2020).
- Nagraj, V. P., Sinkin, R. A., Lake, D. E., Randall Moorman, J. & Fairchild, K. D. Recovery from bradycardia and desaturation events at 32 weeks corrected age and NICU length of stay: an indicator of physiologic resilience? *Pediatr. Res.* <https://doi.org/10.1038/s41390-019-0488-3> (2019).
- Fairchild, K. D., Nagraj, V. P., Sullivan, B. A., Moorman, J. R. & Lake, D. E. Oxygen desaturations in the early neonatal period predict development of bronchopulmonary dysplasia. *Pediatr. Res.* <https://doi.org/10.1038/s41390-018-0223-5> (2018).
- Sullivan, B. A. et al. Early pulse oximetry data improves prediction of death and adverse outcomes in a two-center cohort of very low birth weight infants. *Am. J. Perinatol.* <https://doi.org/10.1055/s-0038-1654712> (2018).
- Vesoulis, Z. A. et al. Early hypoxemia burden is strongly associated with severe intracranial hemorrhage in preterm infants. *J. Perinatol.* **39**, 48–53 (2019).
- Searle, S. R., Speed, F. M. & Milliken, G. A. Population marginal means in the linear model: an alternative to least squares means. *Am. Stat.* **34**, 216–221 (1980).
- Fleming, S. et al. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. *Lancet* **377**, 1011–1018 (2011).
- Alonzo, C. J. et al. Heart rate ranges in premature neonates using high resolution physiologic data. *J. Perinatol.* **38**, 1242–1245 (2018).
- Van Marter, L. J. et al. Do clinical markers of barotrauma and oxygen toxicity explain interhospital variation in rates of chronic lung disease? The

- Neonatology Committee for the Developmental Network. *Pediatrics* **105**, 1194–1201 (2000).
26. Fairchild, K. et al. Clinical associations of immature breathing in preterm infants: part 1-central apnea. *Pediatr. Res.* **80**, 21–27 (2016).
  27. March Of Dimes. Infant mortality rates St. Louis city. *PeriStats*. <https://www.marchofdimes.org/Peristats/ViewSubtopic.aspx?reg=29510&top=6&stop=91&lev=1&slev=5&obj=9>.
  28. March Of Dimes. Infant mortality rates Virginia. *PeriStats*. <https://www.marchofdimes.org/Peristats/ViewSubtopic.aspx?reg=51&top=6&stop=91&lev=1&slev=4&obj=9>.
  29. March Of Dimes. Infant mortality rates Manhattan. *PeriStats*. <https://www.marchofdimes.org/Peristats/ViewSubtopic.aspx?reg=36061&dv=rdo&top=6&stop=91&lev=1&slev=6&eny=&sty=&chy=&obj=8>.
  30. Dennergy, P. A. et al. Pre-Vent: the prematurity-related ventilatory control study. *Pediatr. Res.* **85**, 769–776 (2019).
  31. Ransohoff, D. F. & Feinstein, A. R. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N. Engl. J. Med.* **299**, 926–930 (1978).
  32. Reyna, M. A. et al. Early prediction of sepsis from clinical data. *Crit. Care Med.* **48**, 210–217 (2020).