HHS Public Access

Vital Sign Metrics of VLBW Infants in Three NICUs: Implications for Predictive Algorithms

Amanda M. Zimmet[1], **Brynne A. Sullivan**[2], **Karen D. Fairchild**[2], **J. Randall Moorman**[1], **Joseph R. Isler**[3], **Aaron W. Wallman-Stokes**[3], **Rakesh Sahni**[3], **Zachary A. Vesoulis**[4], **Sarah J. Ratcliffe**[5], **Douglas E. Lake**[1,6]

[1]Department of Medicine, University of Virginia, Charlottesville, VA

[2]Department of Pediatrics, University of Virginia, Charlottesville, VA

[3]Department of Pediatrics, Columbia University, New York, NY

[4]Department of Pediatrics, Washington University St. Louis, St. Louis, MO

[5]Department of Public Health Science, University of Virginia, Charlottesville, VA

[6]Sepsis Challenges Response Unit, University of Virginia, Charlottesville, VA

Abstract

Background: Continuous heart rate (HR) and oxygenation (SpO_2) 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 NICUs, we analyzed the mean, standard deviation, skewness, kurtosis, and cross-correlation of electrocardiogram HR, pulse oximeter pulse rate, and SpO_2. 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 1,168 VLBW infants from birth through day 42 (35,238 infant-days). HR and SpO_2 metrics were similar at the three NICUs, with mean HR rising by ~10 beats/minute over the first two weeks and mean SpO_2 remaining stable around 94% over time. The number of bradycardia events was higher at one site, and the duration of desaturations was longer at another site.

Corresponding Author Karen D. Fairchild, MD, kdf2n@virginia.edu, 434-924-5428, University of Virginia School of Medicine, Department of Pediatrics, Box 800386 Hospital Dr., Charlottesville, VA 22908.

Conclusion: Mean HR and SpO_2 were generally similar among VLBW infants at three NICUs from birth through six weeks of age, but bradycardia and desaturation events differed in the first two weeks after birth. This highlights the importance of developing predictive analytics tools at multiple sites.

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 (1). Subtle changes can occur before overt clinical signs of illness, prompting development of early warning systems that alert clinicians to changes in patient status requiring attention (2–4). One example is the finding of abnormal heart rate characteristics of decreased variability and transient repetitive decelerations that sometimes precede clinical presentation of sepsis, necrotizing enterocolitis (NEC), or other infections in very low birth weight (VLBW) preterm infants (5–9).

In a nine-NICU randomized clinical trial of 3003 VLBW infants, display of a heart rate characteristics index, the fold-increase in risk of sepsis being diagnosed in the next day, was associated with a 22% relative decrease in mortality rate (10). Another example of a change in vital signs in preterm infants is the simultaneous fall in heart rate (HR) and oxygen saturation (SpO_2) during neonatal apnea, the familiar bradycardia-desaturation spell (11–13). A measure of this, the maximum cross-correlation of HR and SpO_2, increased prior to diagnosis of sepsis or NEC in a study of 1065 VLBW infants in two NICUs (14).

HR and SpO_2 are affected not only by illness and stress, but also by maturation and by clinical care practices such as mode of respiratory support (15–17). 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 more than 1000 VLBW infants at three large tertiary care NICUs during the first six weeks of hospitalization. We also compared the number of bradycardia and desaturation events and the cross-correlation of HR and SpO_2. 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.

METHODS

We analyzed vital sign data from VLBW infants (1500 grams birth weight) admitted from 2012–2018 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 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 seven days of HR and SpO_2 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 eight second averaging time. During the study period UVA and WUSTL clinicians had a default goal SpO_2 range of 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 (bpm) at UVA and 100 bpm at the other two sites.

HR, PR, and SpO_2 Metrics

We analyzed continuously measured electrocardiogram-derived HR, pulse oximeter-derived pulse rate (PR), and SpO_2. Daily mean, standard deviation, skewness, and kurtosis of HR, PR, and SpO_2 were computed for each infant over the first six weeks after birth. To control for artifact, all values of zero were removed and, for measurements other than mean, values above the 99th percentile were censored to the 99th percentile value.

Bradycardia and desaturation events were quantified using thresholds and definitions we have previously published (18,19). Bradycardia was defined as HR <100 bpm for at least four seconds and desaturation as SpO_2 <80% for at least 10 seconds. Events were joined if they occurred within four or ten seconds 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_2%. We used our own code written in Matlab for the analyses. Data were smoothed using a sliding window of seven days as we have done in prior work (20,21).

Statistics

We assessed for site effects on each metric using daily means from the day of birth through day 42 by n-way ANOVA. 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, 3 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_2%, etc.) (22). 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 site on a particular metric was measured using $\log_{10}$(p-value), *i.e.*, by reporting the number of leading zeros for the p-value.

RESULTS

During the period of study, 3,209 VLBW infants were admitted to the three NICUs with vital sign data recording available, 1,168 of whom had no exclusions and had at least seven days of stored vital sign data available for analysis in the first four 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 (Supplemental Figure S1).

As shown in Figure 1, the mean HR and SpO_2% were similar at the three sites over the six weeks of study. The mean HR rose from about 150 bpm in the first week to about 160 bpm and changed little thereafter. After two weeks of age, there was a small (about four beats per minute) difference in infants' daily mean HR between sites. The daily mean SpO_2 was slightly different (about 1%) between sites in the first two 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 two-fold; 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 three 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 six weeks. The smaller differences in event durations remained similar throughout. The percentages of time spent in bradycardia or desaturation are shown in Supplemental Figure S2. The number of bradycardia and desaturation events are shown split by birthweight in Supplemental Figure S3.

Although the absolute differences in some of the HR and SpO_2 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 Figure 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 p~<0.001). Metrics are ordered from those with the most to the least inter-site differences. Notably, skewness of pulse rate 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 Figure 3 metrics not shown in Figures 1 and 2 are shown in Supplemental Figures S4–8. Supplemental Figure S9 provides a probability density plot for all vital sign metrics in Figure 3.

Using the average value for each infant for all HR, PR, and SpO_2% 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_2% were significantly different (p<0.05) between the sexes (Supplemental Figure S10), but the differences were small (less than one percent difference in mean SpO_2%).

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, heart rate (HR measured by ECG and PR measured by pulse oximeter) and SpO_2, in the first six 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_2, Figure 1 shows that the overall mean HR increased from about 150 bpm in the first week after birth to about 160 bpm from weeks two to six, while the mean SpO_2 of about 94% was consistent over the time period studied. The change in HR over time that we show in this VLBW cohort is similar to that previously published for term infants, but with an offset of approximately 20 bpm (preterm infants have higher HR than term infants) (23). The HR values in Figure 1 are also similar to those previously published in a single site report on preterm infants at UVA (24). There were inter-site differences of about four bpm in HR and one percent in SpO_2 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 Figure 2, we note that infants at CU had more bradycardia events during the first two weeks after birth. Our definition of bradycardia of <100 bpm was the same as the alarm threshold at CU and WUSTL (and only ten bpm higher than the alarm threshold at UVA) and thus it is unlikely that the difference in 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 (25) leading to more apnea-associated HR decelerations (26). This is also supported by higher cross-correlation of HR and SpO_2 (Supplemental 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_2 between sites, especially in the first week after birth (Supplemental Figure S7) may be an important finding since we identified its association with apnea and exaggerated periodic

breathing (27). Moreover, the cross-correlation of HR and SpO_2 was a significant predictor in a model targeting imminent septicemia or necrotizing enterocolitis (14). In that study of more than 1000 VLBW infants, we also found that infants at CU had a slightly higher baseline cross-correlation of HR and SpO_2 than infants at UVA. The mechanism is unknown, but may relate to less mechanical ventilation at CU and thus more apnea, with concurrent decline in HR and SpO_2.

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, though 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, about one 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 (28–30). 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 (31).

The differences we see here highlight the importance of multi-center 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 analyzed why diagnostic tests fail (32). They advanced the concept, later called *spectrum bias*, that a test is limited if developed on diseased patients that 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 (33). No model that had 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 heart rate 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 (10). Moreover, in more recent work, a HR and SpO_2 model for predicting sepsis performed well at both UVA and CU in spite of differences in vital sign trends (14).

CONCLUSION

In a three-NICU study of 1,168 VLBW infants from birth through six weeks of age, we found that mean HR and SpO_2 were generally similar, but bradycardia and desaturation

events differed in the first two 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

DATA AVAILABILITY

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

REFERENCES

1. 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. (2019) doi:10.1038/s41390-019-0527-0.

2. 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. Informatics 24, 681–692 (2020).

3. Dewhurst CJ, Cooke RWI & Turner MA Clinician observation of physiological trend monitoring to identify late-onset sepsis in preterm infants. Acta Paediatr. Int. J. Paediatr. 97, 1187–1191 (2008).

4. Mithal LB 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). [PubMed: 29351876]

5. Sullivan BA & Fairchild KD Predictive monitoring for sepsis and necrotizing enterocolitis to prevent shock. Semin. Fetal Neonatal Med. 20, 255–61 (2015). [PubMed: 25823938]

6. Griffin MP, Moorman JR, 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). [PubMed: 11134441]

7. Griffin MP et al. Abnormal Heart Rate Characteristics Preceding Neonatal Sepsis and Sepsis-Like Illness. Pediatr. Res. 53, 920–926 (2003). [PubMed: 12646726]

8. Griffin MP et al. Heart rate characteristics: novel physiomarkers to predict neonatal infection and death. Pediatrics 116, 1070–4 (2005). [PubMed: 16263991]

9. Weitkamp JH 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). [PubMed: 31801155]

10. Moorman JRet al.Mortality Reduction by Heart Rate Characteristic Monitoring in Very Low Birth Weight Neonates: A Randomized Trial. J. Pediatr. 159, 900–907 (2011). [PubMed: 21864846]

11. Lee Het al.A new algorithm for detecting central apnea in neonates. Physiol. Meas. 33, 1–17 (2012).

12. Clark MTet al.Breath-by-breath analysis of cardiorespiratory interaction for quantifying developmental maturity in premature infants. J. Appl. Physiol. 112, 859–867 (2012). [PubMed: 22174403]

13. Fairchild KD & Lake DE 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). [PubMed: 29141263]

14. Fairchild KDet 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). [PubMed: 28001143]

15. Martin RJ, DiFiore J. M. Di, Macfarlane PM & Wilson CG Physiologic basis for intermittent hypoxic episodes in preterm infants. Adv. Exp. Med. Biol. 758, 351–8 (2012). [PubMed: 23080182]

16. Bancalari E & Claure N Respiratory Instability and Hypoxemia Episodes in Preterm Infants. Am. J. Perinatol. 35, 534–536 (2018). [PubMed: 29694990]

17. McDonald FB, Dempsey EM & O'Halloran KD The impact of preterm adversity on cardiorespiratory function. Experimental Physiology vol. 105 17–43 (2020). [PubMed: 31626357]

18. Peter Nagraj V, Sinkin RA, Lake DE, Randall Moorman J & Fairchild KD Recovery from bradycardia and desaturation events at 32 weeks corrected age and NICU length of stay: an indicator of physiologic resilience? Pediatr. Res. (2019) doi:10.1038/s41390-019-0488-3.

19. Fairchild KD, Nagraj VP, Sullivan BA, Moorman JR & Lake DE Oxygen desaturations in the early neonatal period predict development of bronchopulmonary dysplasia. Pediatr. Res. (2018) doi:10.1038/s41390-018-0223-5.

20. Sullivan BAet 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. (2018) doi:10.1055/s-0038-1654712.

21. Vesoulis ZAet al.Early hypoxemia burden is strongly associated with severe intracranial hemorrhage in preterm infants. J. Perinatol. 39, 48–53 (2019). [PubMed: 30267001]

22. Searle SR, Speed FM & Milliken GA Population marginal means in the linear model: An alternative to least squares means. Am. Stat. 34, (1980).

23. Fleming Set al.Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: A systematic review of observational studies. Lancet377, 1011–1018 (2011). [PubMed: 21411136]

24. Alonzo CJet al.Heart rate ranges in premature neonates using high resolution physiologic data. J. Perinatol. 38, 1242–1245 (2018). [PubMed: 29925866]

25. Van Marter LJet 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. Pediatrics105, 1194–201 (2000). [PubMed: 10835057]

26. Fairchild Ket al.Clinical associations of immature breathing in preterm infants: part 1-central apnea. Pediatr. Res. 80, 21–7 (2016). [PubMed: 26959485]

27. Fairchild KD & Lake DE 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). [PubMed: 29141263]

28. Infant mortality rates St Louis city | PeriStats | March Of Dimes. https://www.marchofdimes.org/Peristats/ViewSubtopic.aspx?reg=29510&top=6&stop=91&lev=1&slev=5&obj=9.

29. Infant mortality rates Virginia | PeriStats | March Of Dimes. https://www.marchofdimes.org/Peristats/ViewSubtopic.aspx?reg=51&top=6&stop=91&lev=1&slev=4&obj=9.

30. Infant mortality rates Manhattan | PeriStats | March Of Dimes. https://www.marchofdimes.org/Peristats/ViewSubtopic.aspx?reg=36061&dv=rdo&top=6&stop=91&lev=1&slev=6&eny=&sty=&chy=&obj=8.

31. Dennery PAet al.Pre-Vent: the prematurity-related ventilatory control study. Pediatr. Res. 85, 769–776 (2019). [PubMed: 30733614]

32. Ransohoff DF & Feinstein AR Problems of Spectrum and Bias in Evaluating the Efficacy of Diagnostic Tests. N. Engl. J. Med. 299, 926–930 (1978). [PubMed: 692598]

33. Reyna MAet al.Early Prediction of Sepsis From Clinical Data. Crit. Care Med. 48, 210–217 (2020). [PubMed: 31939789]

Impact Statement

Key message:

HR and SpO_2 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.

What does this study add to existing literature:

Although mean HR and SpO_2 patterns were similar in VLBW infants at 3 NICUs, inter-site differences in number of bradycardia events and duration of desaturation events were found.

Impact:

Inter-site differences in bradycardia and desaturation events among VLBW infants should be considered in development of predictive algorithms.
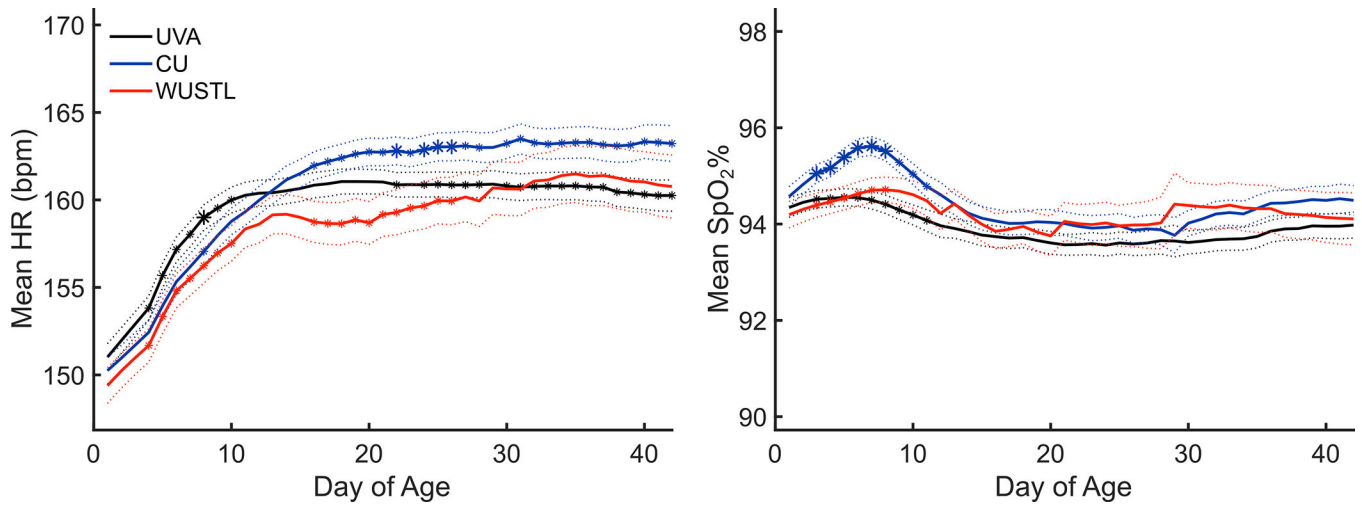
Figure 1: HR and SpO$_2$ Trends of VLBWs at Three NICUs.

Daily mean HR (left) and SpO$_2$ (right) are shown for VLBW infants at the three NICUs through the first six 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.
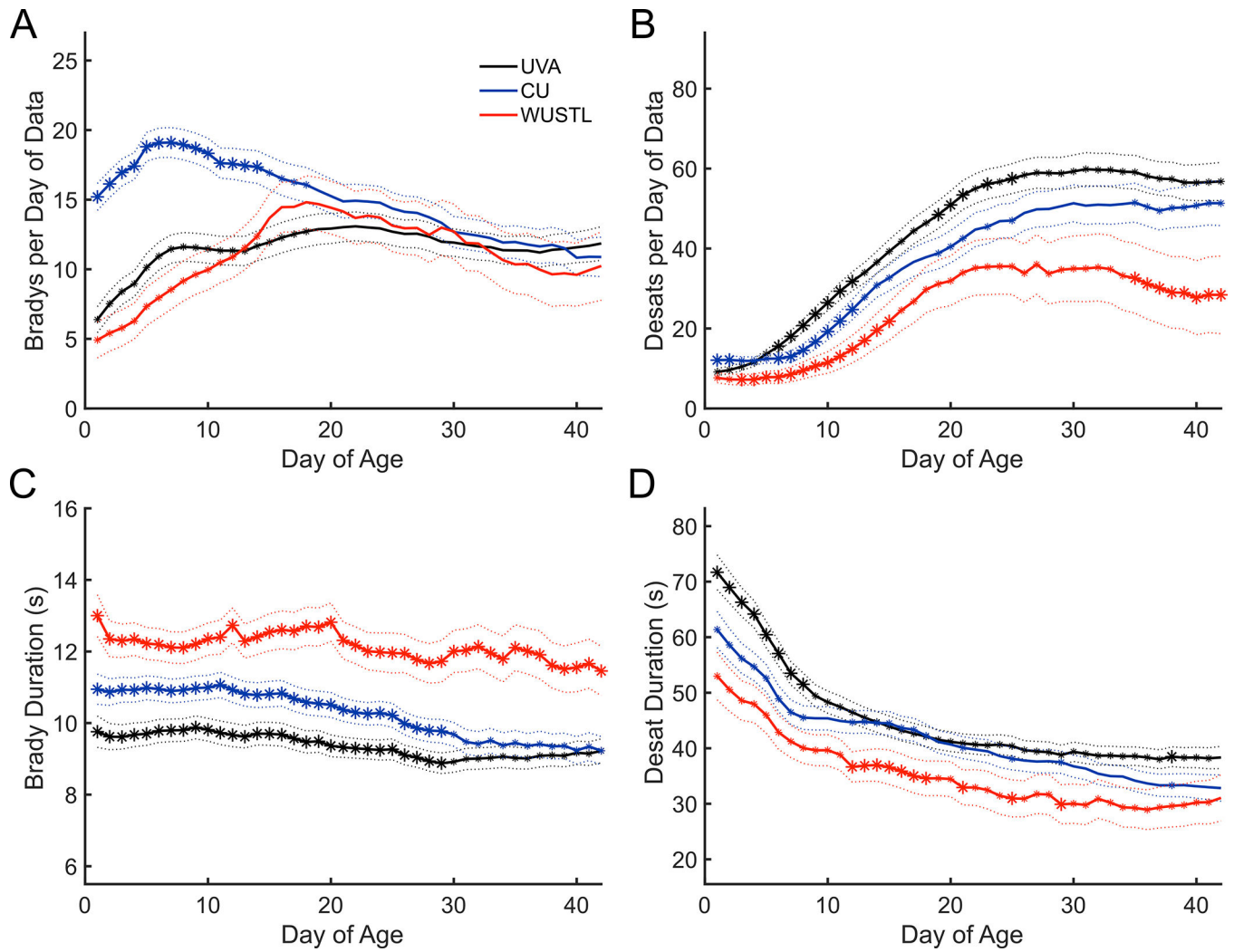
Figure 2: Bradycardia and Desaturation Events by Site.
Mean number of bradycardia (A) and desaturation (B) events per day of data are shown for the first six weeks from birth for VLBW infants at UVA, CU, and WUSTL. Mean event duration in seconds is shown in the bottom panels (C,D). Dotted lines indicate 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.
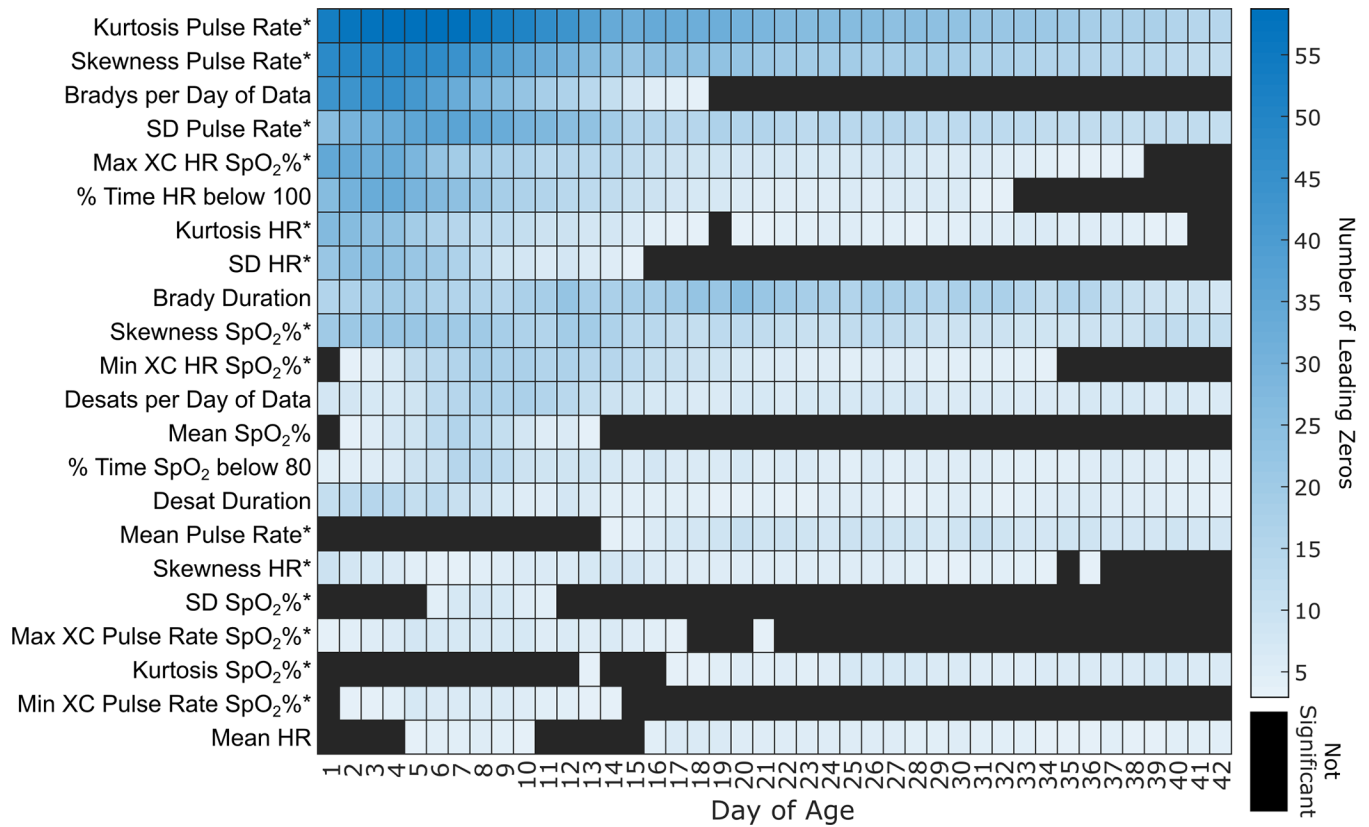
Figure 3: Magnitude of Statistical Significance of Site Differences in HR, PR, and SpO₂ metrics.
For each metric shown on the left y axis, the number of zeros preceding the p-value for
inter-site differences each day shown on the x axis is depicted as a heat map. Black boxes
indicate no significant difference 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 ten minute averages. SD = standard
deviation, XC = cross-correlation, HR = ECG heart rate.

Table 1:

Demographic and Clinical Variables

	UVA	CU	WUSTL
N	477	444	247
Median birth weight, grams (IQR)	1,030 (798–1280)	1,016 (763–1285)	935 (718–1,150)
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	5,796
Median length of stay, days (IQR)	61 (39–96)	63 (41–95)	81 (56–117)
Mortality	3%	4%	15%