CLINICAL RESEARCH ARTICLE

Predictive analytics in the pediatric intensive care unit for early identification of sepsis: capturing the context of age

Michael C. Spaeder^{1,2}, J. Randall Moorman^{2,3,4,5,6}, Christine A. Tran⁷, Jessica Keim-Malpass^{2,8}, Jenna V. Zschaebitz¹, Douglas E. Lake^{2,4,9} and Matthew T. Clark^{2,3}

BACKGROUND: Early recognition of patients at risk for sepsis is paramount to improve clinical outcomes. We hypothesized that subtle signatures of illness are present in physiological and biochemical time series of pediatric-intensive care unit (PICU) patients in the early stages of sepsis.

METHODS: We developed multivariate models in a retrospective observational cohort to predict the clinical diagnosis of sepsis in children. We focused on age as a predictor and asked whether random forest models, with their potential for multiple cut points, had better performance than logistic regression.

RESULTS: One thousand seven hundred and eleven admissions for 1425 patients admitted to a mixed cardiac and medical/surgical PICU were included. We identified, through individual chart review, 187 sepsis diagnoses that were not within 14 days of a prior sepsis diagnosis. Multivariate models predicted sepsis in the next 24 h: cross-validated C-statistic for logistic regression and random forest were 0.74 (95% confidence interval (CI): 0.71–0.77) and 0.76 (95% CI: 0.73–0.79), respectively.

CONCLUSIONS: Statistical models based on physiological and biochemical data already available in the PICU identify high-risk patients up to 24 h prior to the clinical diagnosis of sepsis. The random forest model was superior to logistic regression in capturing the context of age.

Pediatric Research (2019) 86:655–661; <https://doi.org/10.1038/s41390-019-0518-1>

INTRODUCTION

Sepsis remains a major cause of morbidity and mortality in children.^{1–3} Among infants and children in pediatric-intensive care unit (PICU) worldwide, Weiss et al.¹ found the point prevalence for severe sepsis to be 8.2% with an associated hospital mortality of 25%. Early recognition of patients at risk for sepsis to initiate goal-directed therapy is paramount to improve clinical outcomes.

Predictive analytics monitoring holds promise for early detection of subacute potentially catastrophic illness. This new field uses physiologic and biochemical data to develop algorithms to identify signatures of illness present early in the course of decompensation.^{4,5} Scores or indices derived from these algorithms can be displayed to alert the care providers to patients at risk for clinical decompensation. In a multicenter randomized trial of more than 3000 very low birth weight (VLBW) infants, display of a sepsis risk score based on heart rate characteristic (HRCs) analysis reduced mortality by 21%, with a number needed to monitor of 48 infants to prevent one death.⁶ In adult and pediatric patients, predictive monitoring analytics can identify patients at high risk of subacute potentially catastrophic illnesses, such as sepsis,

urgent unplanned intubation, acute hemorrhage, and urgent unplanned ICU transfer up to 24 h in advance.^{7–13}

Age is the primary driver of the profound degree of variability seen in the physiologic data of patients in the PICU.¹⁴ A care provider entering the room of a PICU patient with a heart rate of 50 beats per min will respond quite differently if the patient is 16 days old as opposed to 16 years old. Contextualizing the impact of age on physiology and pathophysiology is imperative to predicting outcomes in the PICU.

We hypothesized that subtle signatures of illness are present in physiological and biochemical time series of PICU patients in the early stages of sepsis. Here, we tested this hypothesis by developing statistical models to identify patients at increased risk for sepsis, and we focused on the impact of age. We hypothesized that the random forest model, with its multiple decision trees allowing fine degrees of risk categorization based on differing age thresholds, would be more effective in capturing age-contextualized pathophysiology than logistic regression with splines. We evaluated our novel risk marker models in the context of established risk markers for pediatric sepsis.¹⁵

¹Department of Pediatrics, Division of Pediatric Critical Care, University of Virginia School of Medicine, Charlottesville, VA, USA; ²Center for Advanced Medical Analytics, University of Virginia, Charlottesville, VA, USA; ³Advanced Medical Predictive Devices, Diagnostics and Displays Inc., Charlottesville, VA, USA; ⁴Department of Medicine, Division of Cardiovascular Medicine, University of Virginia School of Medicine, Charlottesville, VA, USA; ⁵Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, USA; ⁶Department of Molecular Physiology, University of Virginia, Charlottesville, VA, USA; ⁷University of Virginia School of Medicine, Charlottesville, VA, USA; ⁸University of Virginia School of Nursing, Charlottesville, VA, USA and ⁹Department of Statistics, University of Virginia, Charlottesville, VA, USA

Correspondence: Michael C. Spaeder (ms7uw@virginia.edu)

All work pertinent to this manuscript was performed at the University of Virginia Children's Hospital and the University of Virginia School of Medicine, Charlottesville, VA.

Received: 20 December 2018 Revised: 18 April 2019 Accepted: 22 July 2019

Published online: 31 July 2019

MATERIALS AND METHODS

Study design and definitions

The Institutional Review Board at the University of Virginia approved this study. We conducted a retrospective cohort study from December 2013 to May 2016 at the University of Virginia Children's Hospital, an academic, tertiary-care center. The cohort included all admissions to the PICU, a 17-bed combined cardiac and medical/surgical unit with stored continuous physiologic monitoring data. Admission was defined as a unique hospitalization that included PICU stay irrespective of the number of transitions into and out of the PICU. We excluded all data collected while patients were receiving extracorporeal life support.

We established our primary end point of a sepsis event using the 2005 International Pediatric Sepsis Consensus Conference criteria: (1) presence of systemic inflammatory response syndrome (SIRS) and (2) suspected or proven invasive infection caused by any pathogen, including administration of parenteral antibiotics and acquisition of blood culture.¹⁶ Sepsis events were identified by individual review of every chart by a clinician for patients with a blood culture. The time of the sepsis event was established as the time of clinical diagnosis of sepsis (i.e., the earlier of the time of blood culture order or collection), given that the patient met criteria for SIRS in the preceding 12 h and was administered parenteral antibiotics in the subsequent 6 h. To test our hypothesis that signatures of illness precede clinical diagnosis of sepsis, we identified the 12-h window before each event as case data. Data collected at all other times in patients with or without sepsis events were designated as control data. Admissions without archived physiologic monitoring data due to technical complications were excluded. We censored all data in the 14 days following any sepsis event (including sepsis on admission) to allow for resolution of sepsis.

Physiologic data acquisition and predictors

Continuous cardiorespiratory monitoring consisted of waveforms (three leads of electrocardiogram sampled at 240 Hz, pulse plethysmography at 120 Hz, and invasive blood pressure tracings at 120 Hz) and vital signs (heart rate, respiratory rate (RR), peripheral oxygen saturation, invasive blood pressure, ventilator measured RR, and sample-and-hold non-invasive blood pressure) sampled at 0.5 Hz. The GE monitor (GE Healthcare, Chicago, IL) reported sample-and-hold non-invasive blood pressure at 0.5 Hz, while monitor-smoothed invasive blood pressure was measured continuously. We calculated, in 30 min windows with 50% overlap, the following 16 measures: the mean and standard deviation of heart rate (HR), RR, pulse oximetry (SO₂); mean non-invasive systolic and diastolic blood pressure, or invasive blood pressure in its absence (systolic blood pressure and diastolic blood pressure); the three pairwise cross-correlations between HR, RR, and SO₂; standard deviation of heart inter-beat intervals (sRRI); local dynamics score (LDs) and local dynamics density (LDd) of heart inter-beat intervals;¹⁷ coefficient of sample entropy (COSEn);¹⁷ the slope of log variance versus log scale between scales 4 and 12 for detrended fluctuation analysis (DFA) of heart inter-beat intervals.¹⁸ Non-invasive blood pressure was cycled more frequently than q30 min in 95% of epochs in which it was available. Low-quality electrocardiogram data were excluded from cardiac dynamic calculations.^{19,20} LDs and LDd quantify how many heart inter-beat intervals match very many or no other intervals. COSEn quantifies the repeatability of the heart inter-beat interval time series. Finally, DFA quantifies the way in which the variability of heart inter-beat intervals depends upon time scale. The cardiorespiratory dynamics measured from continuous cardiorespiratory monitoring were calculated as described in ref. ⁸ using CoMET® (AMP3D Inc., Charlottesville, VA). Cardiorespiratory dynamics were not available to the care team.

In addition, we extracted oxygen saturation, temperature, Glasgow coma scale, and fraction of inspired oxygen from flowsheets (pulse rate, RR, and blood pressure automatically posted to the electronic

health record from the continuous bedside monitors without reliable validation). We also extracted 11 frequently measured laboratory measurements (serum sodium, potassium, chloride, bicarbonate, blood urea nitrogen, creatinine, glucose, calcium, white blood cell count, hematocrit, platelet count) from the electronic data warehouse. We also included BUN-to-creatinine ratio as a feature. These intermittent features were combined with continuously measured features using sample and hold, and censored when the value was older than 24 and 48 h for vital signs and labs, respectively.

Finally, we included four clinical covariates as features in an effort to capture clinical circumstances that might augment patient risk. We included age as a continuous predictor. We included male gender, the presence of an arterial line, and the presence of mechanical ventilation as binary predictors.

Model development

We developed a model on the entire cohort and used cross-validation performance characteristics. Modeling was performed in R (R Foundation for Statistical Computing, Vienna, Austria) using the *rms* and *randomForest* packages.^{21–23} We used logistic regression and random forest to construct multivariate models that account for the high-dimensional relationships between predictors, and especially to account for the known age dependence of many features. Candidate features were: 16 measures from continuous monitoring, 4 vital signs, 12 laboratory measurements, and the 4 clinical covariates.

We constructed a binary logistic regression model to identify data prior to the sepsis event. Missing data were imputed with median values. The rates of missingness for each 15-min epoch were: moments of continuous vital signs, 1–4%; continuous vital sign cross-correlations, 5–10%; cardiac inter-beat entropy measures, 1%; flowsheet vital signs, 2–4%; basic metabolic panel, 14%; complete blood count, 18.8%. We then built a model and adjusted the covariance matrix for repeated measures from each patient,²⁴ and removed features using fast backward elimination.²⁵ Another model was built with restricted cubic spline transformations using 3 knots of nonlinearity on each feature with enough unique values.²⁶ The model (2 knots or 3) with the highest Akaike Information Criterion was retained. We divided the output of the selected model (probability of sepsis in the next 12 h) by the average rate of sepsis to get relative risk.

We also constructed a random forest to identify data for patients who had a sepsis diagnosis in the next 12 h. The forest consisted of 400 classification trees, and 6 features (the square root of the number of features)²⁷ were sampled as candidates at each split. We adjusted for the imbalance between the sepsis and non-sepsis data by bootstrapping the event data and sampling the same number of non-event data. We also input a white noise feature as a baseline measure of feature importance. The output of the model was the fraction of trees that classified a record as an event. We divided this output by the average value to obtain predicted relative risk, and then multiplied by the average probability of sepsis in the next 12 h to get predicted probability.

For each model the predicted risk was calculated using leave-one-out cross-validation.²⁸ Briefly, the first hospital admission was identified as index admission, and we built a model on the remaining *N*-1 admissions. The predicted risk for the index admission was estimated using this model, and the procedure repeated for each of *N* admissions. In this way risk estimate for each admission were out of sample.

For each model feature importance was determined by calculating the increase in residual sum of squares obtained by permuting the feature. For each model we calculated the residual sum of squares. We then resampled each feature without replacement 30 times and calculated the average increase in residual sum of squares using the permuted feature. This allows direct comparison of feature importance between the models.

Model development: standard risk markers

We also followed the model development and validation procedures described above to construct a risk marker to use as a reference predictor. This standard risk marker was built using the number of SIRS criteria met, and binary features for times when a patient meets each SIRS criterion. We again divided the model output by the average risk for sepsis in the next 12 h in order to get relative risk. The features in the standard risk marker, when combined with clinical context, were used to retrospectively identify the sepsis events and one might reasonably ask how well these risk markers could be used to identify sepsis in real time. We therefore evaluated our models in context of these standard risk markers.

Performance characteristics of predictive models

The models developed here are intended for two use cases: (1) continuous risk estimation and (2) sepsis screening alerts. We contextualized the performance and potential clinical impact of our predictive models relative to the HRC index. Finally, we evaluated our models relative to a standard risk marker model that included the SIRS criteria met by a patient at any given time.

For the use case of continuous risk estimation, we evaluated these models by plotting the time course of predicted risk leading up to the time of clinical sepsis diagnosis. Early detection requires not only high-predicted risk before diagnosis but also an increasing risk to indicate worsening patient status towards critical illness. We also calculated the area under the receiver operating characteristic (AUC) with confidence intervals based on 200 bootstrap runs resampled by admission. Finally, we calculated the net reclassification improvement (NRI) between our models and ensemble models containing our model output and the standard risk marker model. The NRI quantifies the percentage of risk estimates for cases and controls that change in the correct direction (higher risk or lower risk, respectively) when using additional information.

We evaluated the performance of our models when used for alerting, focusing on alerts that sound only infrequently. For this purpose, we selected a range of thresholds and, for each threshold, defined alerts as starting at upward threshold crossings and ending at downward threshold crossings. We then calculated the number of alerts per week, and the positive predictive value (PPV) and sensitivity when an alert was required to start within 24 h prior to clinical diagnosis of sepsis. We excluded alerts that started within 14 days following a sepsis event since such alerts are true positives for sepsis but false positives for early detection.

RESULTS

Study patients and outcomes

There were 1842 admissions involving 1521 patients during the period of study. Table 1 shows demographics of the study populations. One hundred and eighty-seven (10.2%) admissions were associated with a sepsis event: 92 (49.2%) had only one event, 37 (19.8%) had 2 events, 19 (10.2%) had 3 events, 14 (7.5%) had 4 events, and 25 (13.3%) had 5 or more events. Patients with a sepsis event were younger (median age on admission 17 vs. 40 months, $p = 0.0001$). Patients with a sepsis event had a 7.5-fold higher mortality (18.2 vs. 2.4%, $p < 0.0001$), 24 more days of hospitalization ($p < 0.0001$), and 5.6 more days of mechanical ventilation ($p < 0.0001$). In total, there were 511 sepsis events in 166 patients.

Archived physiologic monitoring data were available for 1711 (93%) admissions (1425 patients) and included in the model construction. After accounting for repeat sepsis events in the 14-day interval following an event (including sepsis diagnosis on or before ICU admission), we identified 187 sepsis events occurring over 154 admissions in a total of 136 patients. The most common diagnostic categories for patients with sepsis were cardiac (46%), respiratory (20%), oncologic (7%), neurological (7%), and trauma (7%).

Table 1. Demographic characteristics of the study population

Characteristic	Non-sepsis	Sepsis	<i>p</i> Value
Admissions	1655	187	
Male sex	53.7%	50.3%	0.3702
Age (years)	3.2 (IQR 0.5–11.9)	1.4 (IQR 0.3–7.4)	0.0001
Race			
Asian	14 (0.8%)	4 (2.1%)	0.0884
Black	292 (17.6%)	36 (19.3%)	0.5859
Other	122 (7.4%)	9 (4.8%)	0.1969
Unknown	6 (0.4%)	0 (0.0%)	0.4095
White	1221 (73.8%)	138 (73.8%)	0.9952
Ventilator days	0.0 (IQR 0.0–0.4)	5.6 (IQR 0.7–19.7)	<0.0001
Hospital stay (days)	4 (IQR 2–8)	28 (IQR 9–62)	<0.0001
In-hospital mortality	2.4%	18.2%	<0.0001

IQR interquartile range

Model development

Figure 1 shows the features in the logistic regression (a) and random forest (b) pediatric sepsis prediction models ordered by increase in residual sum of squares. Only the top 9 predictors are shown for random forest for comparison with logistic regression: all 36 features were used in the random forest model, and none were unimportant (increase in residual sum of squares less than white noise). The complete list of random forest feature importance is shown in the Supplementary Appendix. The major components of the models were temperature, HR from continuous monitoring, age, platelet count, and white blood cell count. It is of note that three of the four SIRS criteria used to identify the events were included in both models and comprise three of the five best predictors. Random forest identifies age as the most important predictor, whereas logistic regression ranks it third and much less important than HR and temperature.

Figure 2 shows the univariate risk profile of each feature in the logistic regression (a) and random forest (b) models. As before, only the top 9 predictors in the random forest model are shown for comparison with logistic regression. The log-odds of sepsis is on the ordinate and the measured variable is on the abscissa. For example, high or low white blood cell count predicted higher risk for sepsis.

Note that the random forest model obtains the risk profiles directly and empirically from the data while logistic regression uses a spline function to fit the relationship. Predictor variables that alter risk over only narrow ranges might lose their power if represented by a smoothed function. Here, we were particularly interested in the differences in the risk profiles for age. Logistic regression predicted highest risk at 5 years old and lowest risk for the youngest and oldest patients, changing smoothly in between. Random forest also predicted lowest risk for the very youngest and oldest patients, but found a much steeper rise in risk in the 1 year old, a relationship that not only better captures the large risk changes over narrow age ranges but also resonates better with clinical experience.

Performance characteristics of predictive models: continuous risk estimation

Figure 3 shows the average time course of the out-of-sample estimated risk based on logistic regression (a) and random forest (b) models leading up to the time of sepsis. These plots show, on average, the risk predicted by the models in the days leading up to clinical recognition of sepsis, cf. refs. ^{29,30} The model output is shown as risk relative to the average risk of sepsis in the study population (1.2% for sepsis in the next 12 h). A relative risk of 1.0

Fig. 1 Components of the logistic regression (a) and random forest (b) sepsis prediction models. Features are ordered by increase in residual sum of squares introduced by permuting the feature. Only the top 9 features are shown for random forest for comparison with the logistic regression model. HR, heart rate; Temp, temperature; WBC, white blood cell count; Plt, platelet count; RRxSO₂, cross-correlation of respiratory rate and oxygen saturation; sRRI, standard deviation of cardiac inter-beat intervals; O₂V, oxygen saturation variability; Hct, hematocrit; BUN/sCr, blood urea nitrogen to creatinine ratio; LDd, local dynamics density score for heart inter-beat intervals

indicates the average risk, 2.0-fold indicates twice the average risk, and so on. For both models the average predicted risk for patients who became septic was higher than the average predicted risk for all patients beginning 36 h prior to sepsis diagnosis. The predicted risk from both models rose during the 30 h prior to sepsis and became significantly higher 24 h prior to the diagnosis of sepsis. The relative risk predicted by logistic regression doubled over the 30 h prior to sepsis, from 1.9-fold to 5.3-fold, and predicted risk from random forest increased by 47%, from 1.5-fold to 2.2-fold. Following blood culture, the estimated risk fell, although at a much slower rate than the rise before sepsis. These dynamic

Fig. 2 Risk profile for features in the logistic regression (a) and random forest (b) models. Log-odds of sepsis is shown on the ordinate and the measured feature is on the abscissa. The 95% confidence interval is shown as a gray ribbon

changes in advance of clinical sepsis diagnosis are comparable to that seen in the HRC randomized trial.³⁰

Figure 4 shows the cross-validated AUC of the logistic regression (circles) and random forest (triangles) models for detection of sepsis. Figure 4a shows the AUC as a function of window prior to sepsis that is considered the event. For a detection window of 24 h (i.e., in the day preceding diagnosis), the logistic regression AUC was 0.735 (95% confidence interval (CI): 0.710–0.771) and the random forest AUC was 0.762 (95% CI: 0.728–0.789). Figure 4b shows the AUC for patients of different ages. Results are shown for overlapping quintiles of age at admission. The random forest model has significantly higher AUC for neonates and teenagers than does logistic regression. For comparison, the AUC of HRC monitoring for neonatal sepsis (that resulted in 20% mortality reduction) within 24 h was 0.70 in the development cohort and 0.75 in the validation cohort.³¹

It is important to understand the performance of the models relative to standard risk markers for pediatric sepsis. For a detection window of 24 h the AUC of the standard risk marker model (including number of SIRS criteria and a binary feature for each criterion) was 0.56 (95% CI: 0.52–0.59). Both models added

Fig. 3 Time course of the average model output leading up to the time of sepsis diagnosis. Results are shown for **a** logistic regression and **b** random forest. The 95% confidence interval is shown as a gray band. White points indicate times where the model outputs are significantly higher ($p < 0.05$) than outputs from the same patient 24 h prior. The dashed line shows the mean predicted risk for the entire population

significantly to standard risk markers ($p < 0.0001$). The NRI for adding the logistic regression or random forest model to the standard risk marker model was 0.493 (95% CI: 0.370–0.620) and 0.692 (95% CI: 0.587–0.795), respectively. That is, adding the logistic regression (random forest) model to standard risk markers correctly reclassifies 49.3% (69.2%) of patient risk estimates. The NRI as a function of the change in relative risk required for reclassification is shown in the Supplementary Appendix. For comparison, the NRI for HRC monitoring relative to standard risk markers was 0.389.³²

Figure 5 shows the calibration of the logistic regression (circles) and random forest (triangles) models. Better-calibrated models are closer to the line of identity (dashed). Models with better discrimination have a higher range between highest and lowest

Fig. 4 Area under the receiver operating characteristic (AUC) as a function of **a** the time period before sepsis defined as event, and **b** age at hospital admission. Results are shown for logistic regression (circles) and random forest (triangles). Confidence intervals are based on 200 bootstrap runs resampled by admission. Asterisks identify AUC that are significantly different based on a Wilcoxon's rank-sum test between the 200 bootstrapped AUC

risk. The logistic regression has better discrimination than the random forest model, as expected because of the lower variability of ensemble models such as random forest. The random forest model, however, generally has better calibration than the logistic regression model, especially for the highest risk patients.

Performance characteristics of predictive models: threshold-based alerting

The SIRS criteria are used in the definition of pediatric sepsis, and the number of SIRS criteria are often used for sepsis alerting for hospitalized patients. In a retrospective analysis, patients in our cohort met SIRS criteria 14.3% of the time, and patients at times when SIRS were met had 3.2-fold higher risk for clinical sepsis

Fig. 5 Calibration of logistic regression (circles) and random forest (right) models. Results are shown for a 12-h sepsis detection window before clinical diagnosis. Predicted risk relative to average is on the abscissa and observed risk relative to average is on the ordinate. Each point represents one decile of predicted risk. The line of identity is shown as a dashed line

diagnosis in the next 24 h. Using our logistic regression and random forest models, patients with estimated relative risk >2.0 (9 and 7% of estimates, respectively) had 4.9-fold (95% CI: 4.2–5.7) and 4.4-fold (95% CI: 3.8–5.1) higher risk, respectively, for sepsis than patients with estimated risk <2.0 after adjusting for pediatric SIRS.

We also constructed alerts based on upward and downward threshold crossings. The rate of sepsis is about 1.6 per week. Alerts based on SIRS criteria occurred 23 times per week, lasted 11 h, had a PPV of 6.1% and sensitivity of 62%. Note that the sensitivity was not 100% because this analysis required the SIRS trigger to go from off to on in the 24 h preceding diagnosis. The SIRS trigger was highly sensitive but had a high false-positive rate. For our models, we targeted a much lower alert rate to reduce false positives at the expense of sensitivity. An alert based on our logistic regression model that occurred 2.5 times per week remained on for 2 h, had a PPV of 20.4%, and a sensitivity of 21.5%. An alert based on our random forest model that occurred 2.5 times per week remained on for 1 h, had a PPV of 16.0%, and a sensitivity of 10.2%. That is, infrequent alerts based on our logistic regression and random forest models may have allowed earlier diagnosis and treatment in 21.3% and 10.2% of sepsis cases, respectively. Alert characteristics are shown for a range of alarm rates in the Supplementary Appendix.

DISCUSSION

We have studied physiological and biochemical dynamics near the clinical diagnosis of sepsis in a PICU. We developed predictive analytic models to identify patients in the hours preceding the time of blood culture. Our major finding is that dynamic data already available for patient care identified patients at risk for sepsis up to 24 h prior to the clinical diagnosis: estimated risk was high 24 h prior to sepsis and rose 0.5- to 2.8-fold as patients became septic.

While the relationships of many predictor variables to sepsis risk are the same in the two modeling strategies, the influence of age

on risk differs. The random forest model more distinctly captures the increased risk of sepsis in the very young PICU patient, which in our unit is dominated by neonates and young infants with congenital heart disease. This particular sub-population is at risk for disordered physiology for a great number of reasons other than sepsis. Decreased model performance in the very young is a limitation of logistic regression, even with restricted cubic splines. We suggest that the random forest model may find particular utility in the PICU because of its agility in finding pockets of highly age-specific pathophysiology that identify patients at high risk of adverse events that have prodromal signatures in the physiological and biochemical data. Moreover, we note that there are other problems in clinical medicine in which this property of the random forest model might be useful in predictive analytics monitoring.

Continuous predictive analytics monitoring relies on the scientific premise that the technology and computational models will detect early signs of physiologically adverse events, such as pediatric sepsis.³³ In order to impact health outcomes, continuous predictive analytic technologies must be implemented in an informed way that allows clinicians to move from reactive to proactive clinical action based on personalized surveillance of physiologic states.³³ Additionally, implementation practices must include ongoing clinician input so that efforts can ensure that real-time monitoring optimizes their already established clinical workflow patterns.³⁴

In order for continuous predictive analytics monitoring to be viewed as a useful means of clinical decision support, several processes are necessary for successful adoption, including: understanding the science behind the algorithm, trusting the data, integrating with the electronic medical record, and optimizing clinical pathways.³⁴ As such, continuous predictive analytics can be integrated into existing hospital information technology infrastructure to allow analytics to be available for use at any point in the continuum of care. These processes were implemented for continuous predictive analytics monitoring based on HRC for sepsis in VLBW infants and resulted in more than a 20% relative reduction in mortality.⁶ The AUC of our models are commensurate with that of HRC monitoring for sepsis within 24 h (0.70 at University of Virginia, 0.75 at a validation site). The current study represents a first step towards real-time, prospective implementation of predictive analytics for early detection of sepsis in the PICU. The next steps along this path, as laid out by the HRC index, are external validation of these predictive models at another site followed by a randomized controlled trial to test the impact of prospective clinical use.

Our study is limited by the single-center observational nature. We built our model using data from a mixed cardiac/medical-surgical PICU, which may limit its applicability broadly to all PICUs and patient populations. Physiological monitoring data included epochs with artifact due to, for example, physical therapy or suctioning. We also did not directly control for the impact of routine medications, such as sedatives and vasoactives, on physiology. Given the recent adoption of Sepsis-3 definitions for adult patients, it is a reasonable assumption that pediatric definitions will be reassessed in the near future, which will require revalidation of our models.

CONCLUSIONS

A predictive model based on age, physiological, and biochemical data accurately identified pediatric ICU patients at high risk for sepsis. Predicted risk rose during the 24 h prior to clinical diagnosis of sepsis as a result of subtle changes in, amongst other predictors, markers of inflammation from the SIRS criteria. The random forest model was superior to the logistic regression in capturing the context of age. Providing the accurate risk estimates produced by this model in real time to clinical personnel at the

point of care may allow for earlier and more targeted intervention and may save lives.

ACKNOWLEDGEMENTS

C.A.T. received funding support from NIAID T35 (T35AI060528).

AUTHOR CONTRIBUTIONS

All authors have met the *Pediatric Research* authorship requirements. Specifically: (1) Substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data: M.C.S., J.R.M., C.A.T., J.V.Z., D.E.L., and M.T.C. (2) Drafting the article or revising it critically for important intellectual content: M.C.S., J.R.M., J.K.-M., D.E.L., and M.T.C. (3) Final approval of the version to be published: M.C.S., J.R.M., C.A.T., J.K.-M., J.V.Z., D.E.L., and M.T.C.

ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41390-019-0518-1>) contains supplementary material, which is available to authorized users.

Competing interests: M.T.C. is employed by and holds equity in Advanced Medical Predictive Devices, Diagnostics, and Displays in Charlottesville, VA, which has licensed technologies from the University of Virginia Licensing and Ventures Group. J.R.M. holds equity and is the Chief Medical Officer of Advanced Medical Predictive Devices, Diagnostics, and Displays in Charlottesville, VA, which has licensed technologies from the University of Virginia Licensing and Ventures Group. D.E.L. is a consultant and holds equity in Medical Predictive Sciences Corporation in Charlottesville, VA. The other authors declare no competing interests.

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

REFERENCES

- Weiss, S. L. et al. Global epidemiology of pediatric severe sepsis: the sepsis prevalence, outcomes, and therapies study. *Am. J. Respir. Crit. Care Med.* **191**, 1147–1157 (2015).
- Schlapbach, L. J. et al. Mortality related to invasive infections, sepsis, and septic shock in critically ill children in Australia and New Zealand, 2002–13: a multi-centre retrospective cohort study. *Lancet Infect. Dis.* **15**, 46–54 (2015).
- Hartman, M. E. et al. Trends in the epidemiology of pediatric severe sepsis. *Pediatr. Crit. Care Med.* **14**, 686–693 (2013).
- Moorman, J. R. et al. Predictive monitoring for early detection of subacute potentially catastrophic illnesses in critical care. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2011**, 5515–5518 (2011).
- Fairchild, K. D. Predictive monitoring for early detection of sepsis in neonatal ICU patients. *Curr. Opin. Pediatr.* **25**, 172–179 (2013).
- Moorman, J. R. et al. Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial. *J. Pediatr.* **159**, 900–906 (2011).
- Moss, T. J. et al. Signatures of subacute potentially catastrophic illness in the ICU: model development and validation. *Crit. Care Med.* **44**, 1639–1648 (2016).
- Moss, T. J. et al. Cardiorespiratory dynamics measured from continuous ECG monitoring improves detection of deterioration in acute care patients: a retrospective cohort study. *PLoS ONE* **12**, e0181448 (2017).
- Mao, Q. et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open* **8**, e017833 (2018).
- Politano, A. D. et al. Predicting the need urgent intubation in a surgical/trauma intensive care unit. *Surgery* **154**, 110–116 (2013).
- Moss, T. J. et al. Heart rate dynamics preceding hemorrhage in the intensive care unit. *J. Electrocardiol.* **48**, 1075–1080 (2015).
- Kamaleswaran, R. et al. Applying artificial intelligence to identify physiologic markers predicting severe sepsis in the PICU. *Pediatr. Crit. Care Med.* **19**, e495–e503 (2018).
- Rusin, C. G. et al. Prediction of imminent, severe deterioration of children with parallel circulations using real-time processing of physiologic data. *J. Thorac. Cardiovasc.* **152**, 171–177 (2016).
- O'Leary, F. et al. Defining normal ranges and centiles for heart and respiratory rates in infants and children: a cross-sectional study of patients attending an Australian tertiary hospital paediatric emergency department. *Arch. Dis. Child* **100**, 733–737 (2015).
- Hlatky, M. A. et al. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation* **119**, 2408–2416 (2009).
- Goldstein, B., Giroir, B. & Randolph, A. and the members of the International Consensus Conference on Pediatric Sepsis. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr. Crit. Care Med.* **6**, 2–8 (2005).
- Lake, D. E. & Moorman, J. R. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am. J. Heart Circ. Physiol.* **300**, H319–H325 (2011).
- Peng, C.-K. et al. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* **5**, 82–87 (1995).
- Clark, M. T. et al. Breath by breath analysis of cardiorespiratory interaction for quantifying developmental maturity in premature infants. *J. Appl. Physiol.* **11**, 859–867 (2012).
- Li, Q., Mark, R. G. & Clifford, G. D. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiol. Meas.* **29**, 15–32 (2008).
- R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2017).
- Harrell, F. E. *rms: Regression Modeling Strategies*. R package version 5.1-0.1. <https://CRAN.R-project.org/package=rms> Accessed 2 July 2018.
- Liaw, A. & Wiener, M. Classification and regression by random Forest. *R. News* **2**, 18–22 (2002).
- White, H. Maximum likelihood estimation of misspecified models. *Econometrica* **50**, 1–25 (1982).
- Lawless, J. F. & Singhal, K. Efficient screening of nonnormal regression models. *Biometrics* **34**, 318–327 (1974).
- Harrell, F. E. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis* (Springer, Berlin, 2001).
- Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
- Keim-Malpass, J. et al. Dynamic data monitoring improves predictive analytics for failed extubation in the ICU. *Physiol. Meas.* **15**, 07005 (2018).
- Griffin, M. P. et al. Heart rate characteristics and clinical signs in neonatal sepsis. *Pediatr. Res.* **61**, 222–227 (2007).
- Fairchild, K. D. et al. Septicemia mortality reduction in neonates in a heart rate characteristics monitoring trial. *Pediatr. Res.* **74**, 5705–5775 (2013).
- Griffin, M. P. et al. Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-like illness. *Pediatr. Res.* **53**, 920–926 (2003).
- Lake, D. E., Fairchild, K. D. & Moorman, J. R. Complex signals bioinformatics: evaluation of heart rate characteristics monitoring as a novel risk marker for neonatal sepsis. *J. Clin. Monit. Comput* **28**, 329–339 (2014).
- Bose, E., Hoffman, L. & Hravnak, M. Monitoring cardiorespiratory instability: current approaches and implications for nursing practice. *Intens. Crit. Care Nurs.* **34**, 73–80 (2016).
- Keim-Malpass, J. et al. Advancing continuous predictive analytics monitoring: moving from implementation to clinical action in a learning health system. *Crit. Care Nurs. Clin. N. Am.* **30**, 273–287 (2018).