Initial Development and External Validation of a Biodigital Rapid Alert to Identify Neuromorbidity A-I Bundle (BRAIN AI) Among Critically Ill Children

Christopher Horvat MD MHA[1], Alicia Au MD MS[1], Harry Hochheiser PhD[2], Tanupat Boonchalermbichien MD[2], Kevin Bui BS[2], Neil Munjal MD MS[3], Mark Wainwright MD PhD[4], Qingyang Li MS MS[4], Robert Clark MD[1]

1. Department of Pediatric Critical Care Medicine, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA | 2. Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA | 3. Department of Pediatrics, University of Wisconsin, Madison, WI | 4. Seattle Children's Hospital, Seattle, WA

Introduction Advances in pediatric intensive care medicine have substantially reduced mortality in recent decades, leading the field to pivot its focus to the preservation of childhood neurodevelopmental potential during critical illness. Novel, brain-derived, serum-based biomarkers of brain injury may serve as harbingers of neurological deterioration before clinical signs or symptoms are clearly apparent at the bedside. Coupling novel biomarkers with advanced, point-of-care analytics leveraging structured, time series data harbored by the electronic record can help create clinical decision support tools capable of aiding clinicians in both obviating patient neurological deterioration and applying appropriate stewardship of the use of innovative, potentially costly brain injury biomarkers (**Figure 1**). We describe the development, initial performance characteristics, and external validation of a point-of-care, interoperable decision support pipeline for early identification of neuromorbidity among critically ill children.

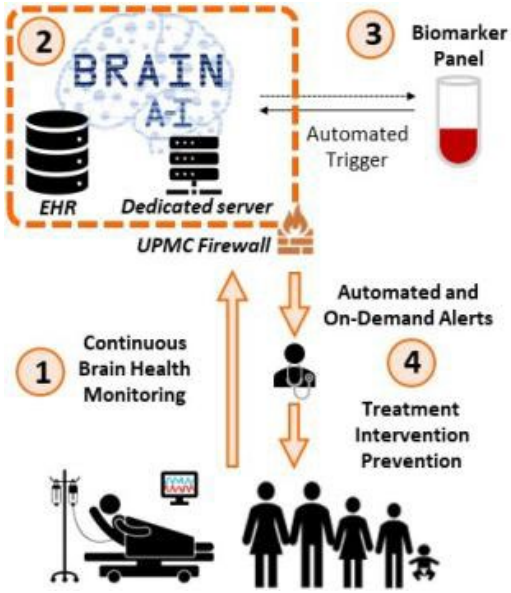

Figure 1 Summary of BRAIN AI

Methods The study centers' institutional review boards approved this work. Initial model development and assessment relied on historic data from a single, freestanding, quaternary children's hospital that admits >2000 patients annually to a 36-bed medical-surgical pediatric intensive care unit (PICU). Data were extracted from an institutional warehouse for all patients admitted to the PICU with a length of stay >24 hours, available demographic data, a recorded nursing physical assessment or laboratory measurement, and an admission diagnosis present for 2009-2018. Candidate data elements were selected based on clinical expertise and with attention to the anticipated United States Core Data for Interoperability (USCDI) requirements. Structured data included age (A), behavior/cognitive assessments (B), surrogates of cardiac arrest (C), drug administrations (D), eye pupil reactivity (E), fluid and electrolyte measurements (F), Glasgow coma scale score (G), vital sign and medications indicating hemodynamic status (H), and vital sign and laboratory indicators of inflammation and invasive support (I), constituting a digital 'A-I bundle.' The outcome of interest was a structured, computable, composite representation of PICU neuromorbidity.[1] The data were divided into a 'full-encounter' dataset leveraging all data until the outcome of interest and a '48-hour' dataset which included data within 48 hours of the outcome of interest, then further divided into train and test sets in a 2:1 ratio. Data were discretized into engineered features by 'vectorizing the patient state' as previously described,[2] which encoded both missingness and temporal information into defined features. Models were developed using

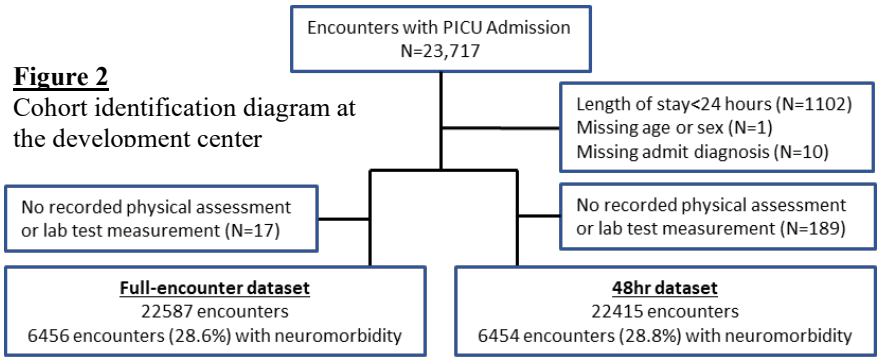

Figure 2
Cohort identification diagram at the development center

correlation-based feature subset selection and information gain feature selection in combination with 5 methods of classification: random forest (RF), logistic regression (LR), Naïve Bayes (NB), support vector machine, and multivariate adaptive regression splines (MARS). Following initial performance assessment, synthetic data representing the necessary structure and schema for model execution were created, bundled with the model software, and shared with a collaborating center for external validation. The validation center is a large, freestanding children's hospital with approximately 2,000 admissions annually to a 39-bed medical-surgical PICU. Neurointensive care consult orders were selected as the outcome of interest for model performance assessment due to differences in data availability at the validation center. Data ingestion, feature engineering, feature selection and classifier approaches were otherwise identical. Levels of the brain-derived serum biomarkers ubiquitin C-terminal hydrolase-L1 (UCH-L1), glial fibrillary acidic protein (GFAP), myelin basic protein (MBP), neuron-specific enolase (NSE), S100 calcium binding protein B (S100B), and spectrin breakdown product 150 (SBDP150) were available for a subset of diagnostically diverse, critically ill children at the development center. Maximum biomarker levels were compared between patients with and without neuromorbidity using the Wilcoxon Rank Sum test (median). Data curation and analyses were conducted using Python (v3.9.7) and R (v4.0.2).

Results From 2009-2018, 22,587 PICU encounters were identified at the development center and were stratified into training and test sets (**Figure 2**). Models constructed using information gain contained 280 features versus 15 features using correlation-based feature selection. The top performing models based on F1-score at the development center used MARS, LR, and RF classification, all with a number needed to alert (NNA) of 2 (**Table 1**). Models demonstrated comparable discrimination at the validation center and lower precision, with an NNA of 3 among top-performing models (**Table 1**). In the subset of 103 diagnostically diverse, critically ill children from the development center, maximum values of UCH-L1, GFAP, NSE, S100B, and SBDP150 were all increased in patients meeting the composite definition of neuromorbidity vs. those that did not (all $P<0.001$).

Table 1 Performance of top-performing models at the development and validation centers

Model	Time Window	Feature Selection	AUROC	AUPRC	F1-score ('True' class)	NNA ('True' class)
Top 3 Development Center Models (Test Years 2016-2018, N=7,007 encounters)						
MARS	Full	Information Gain	0.88	0.73	0.65	2
LR	Full	CFS	0.88	0.73	0.62	2
RF	Full	Information Gain	0.87	0.70	0.61	2
Top 3 Validation Center Models (Test Year 2019, N=1,806 encounters)						
LR	Full	CFS	0.85	0.42	0.40	3
NB	Full	CFS	0.83	0.38	0.40	3
LR	Full	Information Gain	0.81	0.35	0.35	3

Abbreviations: MARS, multivariate adaptive regression splines; LR, logistic regression; RF, random forest; NB, naïve Bayes; CFS, correlation-based feature selection; AUROC, area under the receiver operating characteristics curve; AUPRC, area under the precision recall curve; NNA, number needed to alert (1/positive predictive value)

Discussion A pipeline for training models for predicting neurological deterioration among critically ill children using data elements prioritized by USCDI mandates demonstrated very good performance at both development and external validation centers. Serum brain-derived biomarkers differed significantly between patients meeting vs. not meeting the computable composite definition of neuromorbidity. These results suggest continued effort is warranted to develop, refine and further validate an interoperable decision support tool that leverages available electronic record data and incorporates novel biomarkers of brain injury.

Acknowledgements and Funding We would like to thank Henry Ogoe, PhD, Amie Barda, PhD and Andrew Bauman, PhD for their early involvement in this project. Supported by R01 NS118716. Patent applications U.S. 17/760,558 and PCT/US2020/061985 have been filed related to this work.

References

1. Alcamo AM, Clark RSB, Au AK, Kantawala S, Yablonsky EJ, Sindhi R, et al. Factors associated with neurobehavioral complications in pediatric abdominal organ transplant recipients identified using computable composite cefinitions. Pediatr Crit Care Med J Soc Crit Care Med World Fed Pediatr Intensive Crit Care Soc. 2020 Sep;21(9):804–10.
2. Hauskrecht M, Batal I, Valko M, Visweswaran S, Cooper GF, Clermont G. Outlier detection for patient monitoring and alerting. J Biomed Inform. 2013 Feb 1;46(1):47–55.