I-WIN: an Intensive care Warning INdex System for Early Prediction of Clinical Deterioration and Intervention using Data-driven Machine Learning Modeling of Electronic Health Record Data

Fuchiang (Rich) Tsui, PhD[1-3], Victor Ruiz, PhD[1], Michael Goldsmith, MD[2,3], Lingyun Shi, MS[1], Matthew Scharf[1], Maryam Naim, MD[2,3], and Allan Simpao, MD[2,3]

[1]Tsui Laboratory, [2]Department of Anesthesiology and Critical Care Medicine, Children's Hospital of Philadelphia; [3]Department of Anesthesiology and Critical Care, University of Pennsylvania, PA

ABSTRACT

Given increased challenges in intensive care, there is a critical need to develop predictive models for early warning of patient deterioration. We built and evaluated innovative data-driven machine learning models within the Intensive Care Warning Index (I-WIN) system for early prediction of deterioration and pharmaceutical interventions at a tertiary-care children's hospital. I-WIN currently collects and displays real-time clinical data including EHRs and bedside monitor waveforms. The I-WIN deterioration prediction model with 1,028 variables from various EHR data types (vital signs, laboratory results, diagnoses, and medications) achieved an area under the receiver operating characteristic curve (AUC) of 0.92 (95% CI 0.84-0.98), 0.881 sensitivity, 0.776 positive predictive value, 0.862 specificity, and 0.571 Brier skill score four hours before the onset of clinical deterioration. The pharmaceutical intervention prediction model with 151 variables had AUCs raging from 0.697 (0.58,0.801) to 0.857 (0.767,0.922) for predicting six medications administered in emergency situations. We plan to evaluate the deployment of the I-WIN models following event-driven mechanism and industry technology implementation standards such as Docker and Kubernetes for timeliness, portability, and scalability.

INTRODUCTION

Intensive care units (ICUs) face key challenges and needs despite continuous investment in medical technology and personnel training. Common challenges faced in ICUs include high mortality rates compared with other hospital units (10%-29% in adult ICUs; 2%-6% in pediatric ICUs), increased admissions, medical errors, intensivist shortage, and alert fatigue.[1] Despite state-of-the-art approaches showing promising advantages of using AI in ICUs, current research is limited in several aspects: the use of high dimensional electronic health record (EHR) data, intervention recommendation, and translational science to clinical practice mainly due to infrastructure barriers.[2] In this study, we built and evaluated an AI system, the Intensive Care Warning Index (I-WIN) system[3] at the Children's Hospital of Philadelphia (CHOP) in two clinical aspects: deterioration prediction[4] and intervention prediction.

METHODS

Cohort Data: The CHOP Institutional Review Board (IRB) approved this study (IRB 18-015520). The study period was between 2014 and 2019. The study cohort included infants with palliated or unrepaired single-ventricle, ductal-dependent, or shunt-dependent congenital heart diseases admitted to a cardiac intensive care unit (CICU) at CHOP from birth until one of the following censoring events: age 6 months, conversion to biventricular circulation, Stage 2 palliation, heart transplant, or death. Additional inclusion criteria included required stage-1 single-ventricle palliation, pulmonary-artery banding followed by single-ventricle palliation, surgical systemic to pulmonary shunt, trans-catheter stenting of the ductus arteriosus, or maintenance of ductus arteriosus patency with prostaglandin-E as the primary source of pulmonary or systemic blood flow until the time of a censoring event.[3] Exclusion criteria were admission in the CICU for <24 hours or died within 24 hours of CICU discharge. Cases were defined as a composite of critical deterioration events: emergent intubation, cardiopulmonary resuscitation (CPR), extracorporeal membrane oxygenation (ECMO) cannulation, or CPR with refractory cardiac arrest requiring ECMO (ECPR). Four data types were used in predictive models: vital signs from nursing flowsheets, laboratory test results, administered medications, and coded diagnoses.

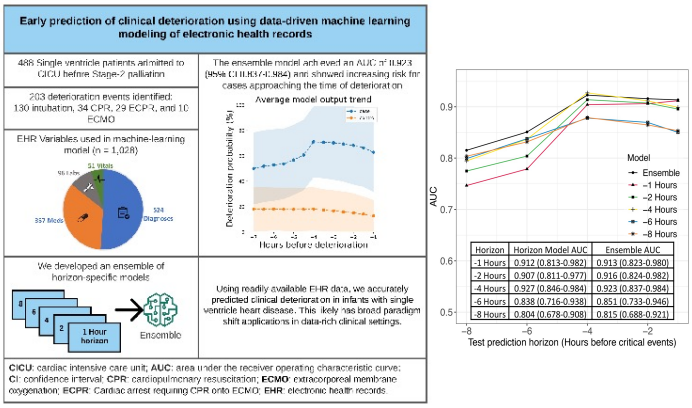
Predictive modeling of deterioration events: For deterioration prediction, we developed an ensemble of 5 machine learning models (1,028 variables) using extreme gradient boosting (XGB) machine algorithm, and each of the models was developed from one of the 5 prediction horizons ranging from 1 hour to 8 hours before the onset of an event.[4]

Predictive modeling of intervention (medications): We used random forest for intervention prediction of medication categories in the same cohort. Six commonly administered medication (intervention) categories used in emergency situations were calcium gluconate, chlorothiazide, fentanyl, furosemide, morphine, and sodium bicarbonate.

Model evaluation We performed a nested 10-fold cross validation for modeling and evaluation. The area under the receiver operating characteristic curve (AUC) was used for evaluation metric.

RESULTS

Deterioration prediction: The cohort included 488 infants with single ventricle physiology with 203 critical events in 134 infants. At 4 hours before deterioration, the model achieved an AUC of 0.92 (95% CI: 0.84-0.98), 0.881 sensitivity, 0.776 positive predictive value, 0.862 specificity, and 0.571 Brier skill score. Performance remained high at 8 hours before deterioration with AUC 0.815 (0.688-0.921).



Horizon	Horizon Model AUC	Ensemble AUC
-1 Hours	0.912 (0.813-0.982)	0.913 (0.823-0.980)
-2 Hours	0.907 (0.811-0.977)	0.916 (0.824-0.982)
-4 Hours	0.927 (0.846-0.984)	0.923 (0.837-0.984)
-6 Hours	0.838 (0.716-0.938)	0.851 (0.733-0.946)
-8 Hours	0.804 (0.678-0.908)	0.815 (0.688-0.921)

Intervention prediction: In the window (8 hours to 30 minutes before deterioration onset), we observed AUCs of 0.857(0.767-0.922) for calcium gluconate, 0.853(0.763-0.923) for fentanyl, 0.741(0.583-0.886) for sodium bicarbonate, 0.718(0.632-0.802) for furosemide, 0.697(0.58-0.801) for morphine, and 0.675(0.511-0.808) for chlorothiazide.

Intervention Prediction Performance

Medicine	ROC AUC	F1	Precision	Recall
Calcium gluconate	0.857 (0.767,0.922)	0.416 (0.17,0.654)	0.421 (0.167,0.698)	0.422 (0.174,0.696)
Chlorothiazide	0.675 (0.511,0.808)	0.177 (0.0,0.356)	0.177 (0.0,0.383)	0.185 (0.0,0.37)
Fentanyl	0.853 (0.763,0.923)	0.446 (0.291,0.637)	0.444 (0.242,0.678)	0.466 (0.267,0.765)
Furosemide	0.718 (0.632,0.802)	0.567 (0.459,0.678)	0.56 (0.426,0.692)	0.579 (0.468,0.692)
Morphine	0.697 (0.58,0.801)	0.192 (0.0,0.346)	0.195 (0.0,0.385)	0.197 (0.0,0.359)
Sodium bicarbonate	0.741 (0.583,0.886)	0.315 (0.1,0.5)	0.312 (0.087,0.579)	0.333 (0.107,0.556)
Weighted average	0.75 (0.692,0.811)	0.425 (0.345,0.512)	0.427 (0.34,0.511)	0.431 (0.344,0.537)

CURRENT STATUS AND PLANS

I-WIN currently is deployed at CHOP as an independent homegrown tool with pilot users from CICU intensivists for patient monitoring and case review. I-WIN uses secure web-based graphical user interfaces and adopts organizational user authentication to enhance usability. I-WIN collects and monitors high-speed waveform and EHR data from over 200 bedside monitors. We plan to add waveform data to the predictive models and embed the models' outputs with explanation to I-WIN following the event-driven approach. IRB approval for conducting user feedback has been obtained. We will evaluate usability, workflow, and the functional effectiveness of the I-WIN system.

REFERENCES

1. Society of Critical Care Medicine. Critical care statistics.

2. Ruiz VM, Saenz L, Lopez-Magallon A, Shields A, Ogoe HA, Suresh S, Munoz R, Tsui FR. Early prediction of critical events for infants with single-ventricle physiology in critical care using routinely collected data. *J Thorac Cardiovasc Surg*. 2019;158(1). doi:10.1016/j.jtcvs.2019.01.130

3. Tsui FR, Shi L, Ruiz V, Mi F, Goldsmith MP, Naim MY, Galvez JA, Simpao AF. Design and deployment of a real-time AI-based telehealth system for deterioration prediction of critical-care patients in a large children's hospital. In: *AMIA*. ; 2020.

4. Ruiz V, Goldsmith M, Shi L, Simpao A, Galvez J, MY N, Nadkarn V, Gaynor W, Tsui F. Early prediction of clinical deterioration using data-driven machine learning modeling of electronic health records. *J Thorac Cardiovasc Surg*. 2021. doi:doi.org/10.1016/j.jtcvs.2021.10.060

CONTACT

Rich Tsui, PhD, FAMIA

tsuif@chop.edu