



Published in final edited form as:

Pediatr Crit Care Med. 2020 August ; 21(8): e572–e575. doi:10.1097/PCC.0000000000002347.

Validation of an Electronic Pediatric Index of Mortality 2 (ePIM2) Score in a Mixed Quaternary Pediatric Intensive Care Unit

Emily L. Joyce, MD, MS^{1,2,3}, Christine M. Crana, MD¹, Jonathan Yabes, PhD^{4,5}, John A. Kellum, MD^{2,3,6}

¹Division of Nephrology, Department of Pediatrics, UPMC Children's Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15261, USA

²Center for Critical Care Nephrology, UPMC and University of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15261, USA

³CRISMA Laboratory (Clinical Research, Investigation, and Systems Modeling of Acute Illness), University of Pittsburgh School of Medicine, Pittsburgh, PA, 15261, USA

⁴Division of General Internal Medicine, Department of Medicine, UPMC, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15261, USA

⁵Department of Biostatistics, University of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15261, USA

⁶Department of Critical Care Medicine, UPMC, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15261, USA

Abstract

Objective: Assess the validity of an electronic version of the Pediatric Index of Mortality 2 (ePIM2) score.

Design: Retrospective observational study.

Setting: Pediatric and cardiac intensive care units (ICUs) at a quaternary medical center.

Patients: Patients older than 60 days of age admitted to the PICU or CICU between 1/1/2010 and 12/31/2014.

Interventions: None

Measurements and Main Results: After adapting the PIM2 score into a version applicable to retrospective electronic health record data, it was validated in a mixed ICU cohort. A manually ascertained PIM2 score was directly compared to the electronically derived ePIM2 score in 100 randomly selected patients with good agreement between score components with nine out of 11 components having an intraclass correlation coefficient (ICC) or Cohen's κ \geq 0.6. In assessing the ePIM2 score in the entire cohort of 12,582 patient encounters, it had good discrimination with

Corresponding Author: Dr. Emily Joyce, 3347 Forbes Ave., Suite 220, Pittsburgh, PA 15213, emilylaurenjoyce@gmail.com.

Article Tweet: Your guide to translating the PIM2 severity of illness scoring system into an electronic health record-friendly version

AUC of 0.89, appropriate calibration with no significant difference between observed and expected deaths, and excellent predictive ability with a Brier score of 0.0135.

Conclusions: The PIM2 score can be adapted to utilize retrospective electronic health record data with acceptable discrimination, calibration and accuracy a large mixed ICU cohort.

Keywords

Mortality; Pediatric Intensive Care Unit; Severity of Illness Index; Pediatric Index of Mortality; Electronic Health Record; Validation Studies

Introduction

Severity of illness scoring is commonly used in intensive care units (ICUs) for both quality assurance and research. The Pediatric Index of Mortality 2 (PIM2) score is one of the most commonly used measures to assess risk of mortality on admission to the ICU (1). Electronic health record (EHR) research is on the rise and effective tools to analyze clinical data is essential for eResearch, which optimizes the integration of information technology and information management (2,3). Research in critically ill children is often confounded by severity of illness, so having a reliable and valid electronic version of the PIM2 score would allow researchers to control for severity of illness. Many previous studies have validated the PIM2 score both prospectively and retrospectively in the pediatric ICU population (4–9). While other commonly-used severity of illness scores in adults such as APACHE and SOFA scores have been adapted to use in EHR research (10,11), PIM2 does not yet have an EHR adaptation.

We assessed each of the 11 data elements that contribute to the PIM2 score and adapted them for applicability within the EHR. An ePIM2 (electronic PIM2) score was then calculated for each patient encounter in a large, granular database in a mixed quaternary ICU setting. Two aims of this study were 1.) to assess the accuracy of the ePIM2 score by comparing it to manually ascertained values from the EHR for a small subset of patients and 2.) to determine the performance of the ePIM2 score for predicting mortality in a large cohort.

Methods

We utilized the Pediatric High-Density Intensive Care (Peds HiDenIC) database, a large academic medical center database of critically ill children generated using EHR data. This database includes patient encounters of children older than 60 days of age, admitted to the pediatric or cardiac ICU at UPMC Children’s Hospital of Pittsburgh between January 2010 and December 2014. The Institutional Review Board waived informed consent. The Peds HiDenIC database had 12,582 patient encounters that qualified for inclusion in analyzing ePIM2 performance. There were initially 12,806 patient encounters, with 80 encounters excluded for missing admission information, 109 encounters excluded for an ICU length of stay less than 4 hours, and 35 encounters excluded for lack of mortality determination.

PIM2 and ePIM2 Agreement

For manual PIM2 ascertainment, 100 randomly selected patients from the database were identified and PIM2 scores were retrospectively developed for each patient encounter utilizing data available in the EHR closest to the time of ICU admission. The Institutional Review Board approved the medical chart review. Two authors, ELJ and CMC, independently collected and confirmed data accuracy. These PIM2 scores were compared to the electronically derived PIM2 (ePIM2) scores for each of the patient encounters utilizing a Bland-Altman Plot. For each data element within the PIM2 score, intraclass correlation coefficient (ICC) or Cohen's κ were developed as appropriate for continuous and categorical variables respectively. Excel (Excel®, Microsoft Corp.) and STATA 15.1 (Stata Corp, TX, USA) were used for data analysis.

ePIM2 Performance

We additionally evaluated the overall predictive ability of the ePIM2 score for each patient encounter in the entire Peds HiDenIC cohort (after exclusions were applied). We used the area under the receiver operating curve (AUC) to assess discrimination and the Hosmer-Lemeshow test and Brier Score to assess calibration. These were performed using STATA 15.1 (Stata Corp, TX, USA).

Results

PIM2 and ePIM2 Agreement

A Bland-Altman plot comparing the difference and average values between mortality risk generated by the two scores for each of the 100 patients is shown in Figure 1 (see Supplementary Figure 1 for Bland-Altman plot comparing the difference and average values of the actual scores). The mean difference between the electronic and manual mortality risk was -0.78% (95% CI -1.68% to 0.12%), and out of the 100 data points, five fell outside the estimated limits of agreement.

Each of the 11 components of the score were compared between the ePIM2 and PIM2 scores. The only two variables that had ICC or κ scores below 0.6 were “elective admission” (κ 0.56, 95% CI 0.37-0.74) and “high risk diagnosis” (κ 0.52, 95% CI 0.24-0.80). The other components were as follows: systolic blood pressure (ICC 0.77, 95% CI 0.68-0.84), pupillary reaction (κ 1), fraction of inspired oxygen (ICC 0.89, 95% CI 0.84-0.82), partial pressure of oxygen in arterial or capillary blood (ICC 1), base excess/deficit (ICC 0.93, 95% CI 0.89-0.95), mechanical ventilation (κ 0.84, 95% CI 0.73-0.96), admitted for surgical recovery (κ 0.65, 95% CI 0.47-0.83), cardiopulmonary bypass (κ 1), and low risk diagnosis (κ 0.68, 95% CI 0.51-0.84).

ePIM2 Performance

The cohort of 12,582 patient encounters was 56.9% male with an average age of 8 years (standard deviation 7.5 years) and average length of ICU admission of 3.9 days (standard deviation 9.6 days). The ePIM2 score predicted 205.8 deaths with 206 deaths actually observed in the cohort (mortality rate of 1.6%), correlating with a standardized mortality ratio of 1.00 (95% CI 0.87-1.15). The AUC was 0.89 (95% CI 0.88-0.89), comparable to the

original PIM2 score AUC of 0.90 (95% CI 0.89-0.91), and the Brier score was good at 0.0135. The population used for the original PIM2 score was younger and had a higher mortality compared to the population used for the ePIM2 score (see Supplemental material for subgroup analyses, comparison to original PIM2 score, and further information on the ePIM2 predictors). The observed and expected mortality using the ePIM2 score over deciles of risk is demonstrated in Table 1, with overall excellent calibration and no significant difference between observed and expected deaths (χ^2 4.19, 8 df, $p=0.84$).

Discussion

The PIM2 scoring system has been used to predict mortality risk, allowing estimation of severity of illness in the pediatric ICU. This is the first study to assess an adapted version of the PIM2 score, which can be applied to retrospective EHR data in critically ill children.

The initial step was to develop the electronic PIM2 (ePIM2) score using the same variables as PIM2. There are 11 primary components of the PIM2 score that are ideally measured at the time of first “face-to-face contact” with the ICU team to minimize lead-time and treatment bias (1). For simplicity, we utilized variables (e.g. vital signs and arterial blood gases) closest to or after the time of ICU admission. A prior study showed no significant difference between PIM2 scores determined between direct and transfer admissions (12), supporting the timing flexibility for use in EHR data. Given the possible limitations of timely documentation and available EHR data in a critically ill patient, we extended the timeframe from assessment within the first one hour of ICU admission to within the first four hours of ICU admission for appropriate variables. Sankar et al showed that interventions within the first 4 hours of ICU admission did not affect the predictive ability of the PIM2 score (13). Each of the 11 primary components were adapted to capture structured data routinely documented and available retrospectively via the EHR (see Supplementary Table 1). An ePIM2 score was then calculated for each ICU encounter in the Peds HiDenIC database using the PIM2 scoring system and adapted variables from EHR. The ePIM2 score performed well, though performed better in the PICU population than in the CICU population (see Supplemental material).

The analysis comparing the agreement between ePIM2 and a manually ascertained PIM2 score showed there was fair agreement between nine out of 11 data elements. The only two variables that had κ scores below 0.6 were “elective admission” and “high risk diagnosis,” which are two of the variables most susceptible to clinician judgement and more difficult to capture with readily available EHR data (e.g. ICD codes). Considering use of more broad ICD code categories may improve sensitivity with these variables. In comparing the overall mortality risk, the bias in the Bland-Altman plot was low at -0.78% , suggesting minimal discrepancy between the manually ascertained and electronically derived mortality risk with 95% of the values falling within the limits of agreement. Overall, our analyses show an acceptable electronic scoring system using a large mixed-ICU cohort with good discrimination, calibration and prediction accuracy though modest discrepancies in granular data (i.e. suboptimal κ values).

There are several limitations in this study. First, the sample size for assessing PIM2 and ePIM2 scores was relatively small with 100 patients. Secondly, the data elements within the PIM2 score generally agree well between manually determined and electronically derived versions, though there is greater variability in less quantitative variables (e.g. elective and surgical admission). Additionally, while the score was applied to a large dataset across two types of ICUs, this remains a single-center study and will require additional validation at other centers and in different EHRs.

Conclusions

While prospective assessment of a PIM2 score is ideal for assessing mortality risk, when unavailable it can be converted into a validated electronic form (ePIM2 score) for utilization as a marker of severity of illness in retrospective EHR research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank the Biostatistical and Data Management Core within the CRISMA (Clinical Research Investigation and Systems Modeling of Acute Illness) Center for their help in variable development and validation as well as data acquisition, management and storage.

Conflicts of Interest and Source of Funding: Drs. Joyce and Crana received support for article research from the National Institutes of Health. The remaining authors have disclosed that they do not have any potential conflicts of interest.

References

1. Slater A, Shann F, Pearson G, et al. PIM2: a revised version of the Paediatric Index of Mortality. *Intensive Care Med.* 2003;29(2):278–285. [PubMed: 12541154]
2. Joyce EL, DeAlmeida DR, Fuhrman DY, et al. eResearch in acute kidney injury: a primer for electronic health record research. *Nephrol Dial Transplant.* 2019;34(3):401–407. [PubMed: 29617846]
3. Kellum JA, DeAlmeida DR, Watzlaf VJ. eResearch: the case of acute kidney injury. *Intensive Care Med.* 2013;39(3):522–523. [PubMed: 23223823]
4. Eulmesekian PG, Perez A, Minces PG, et al. Validation of pediatric index of mortality 2 (PIM2) in a single pediatric intensive care unit of Argentina. *Pediatr Crit Care Med.* 2007;8(1):54–57. [PubMed: 17251883]
5. Netto AL, Muniz VM, Zandonade E, et al. [Performance of the Pediatric Index of Mortality 2 in a pediatric intensive care unit]. *Rev Bras Ter Intensiva.* 2014;26(1):44–50. [PubMed: 24770688]
6. Ng DK, Miu TY, Chiu WK, et al. Validation of Pediatric Index of Mortality 2 in three pediatric intensive care units in Hong Kong. *Indian J Pediatr.* 2011;78(12):1491–1494. [PubMed: 21617907]
7. Wolfler A, Silvani P, Musicco M, et al. Pediatric Index of Mortality 2 score in Italy: a multicenter, prospective, observational study. *Intensive Care Med.* 2007;33(8):1407–1413. [PubMed: 17549454]
8. Imamura T, Nakagawa S, Goldman RD, et al. Validation of pediatric index of mortality 2 (PIM2) in a single pediatric intensive care unit in Japan. *Intensive Care Med.* 2012;38(4):649–654. [PubMed: 22270470]
9. Sankar J, Singh A, Sankar MJ, et al. Pediatric Index of Mortality and PIM2 scores have good calibration in a large cohort of children from a developing country. *Biomed Res Int.* 2014;2014:907871. [PubMed: 25025075]

10. Brundin-Mather R, Soo A, Zuege DJ, et al. Secondary EMR data for quality improvement and research: A comparison of manual and electronic data collection from an integrated critical care electronic medical record system. *J Crit Care.* 2018;47:295–301. [PubMed: 30099330]
11. Huerta LE, Wanderer JP, Ehrenfeld JM, et al. Validation of a Sequential Organ Failure Assessment Score using Electronic Health Record Data. *J Med Syst.* 2018;42(10):199. [PubMed: 30218383]
12. Rahiman S, Sadasivam K, Ridout DA, et al. Comparison of three different timeframes for pediatric index of mortality data collection in transported intensive care admissions*. *Pediatr Crit Care Med.* 2014;15(3):e120–127. [PubMed: 24395001]
13. Sankar J, Chandel A, Dubey NK, et al. Do interventions in an ICU affect the predictive ability of pediatric index of mortality and pediatric index of mortality-2 scores in a tertiary care hospital? *Pediatr Crit Care Med.* 2013;14(2):e70–76. [PubMed: 23287905]

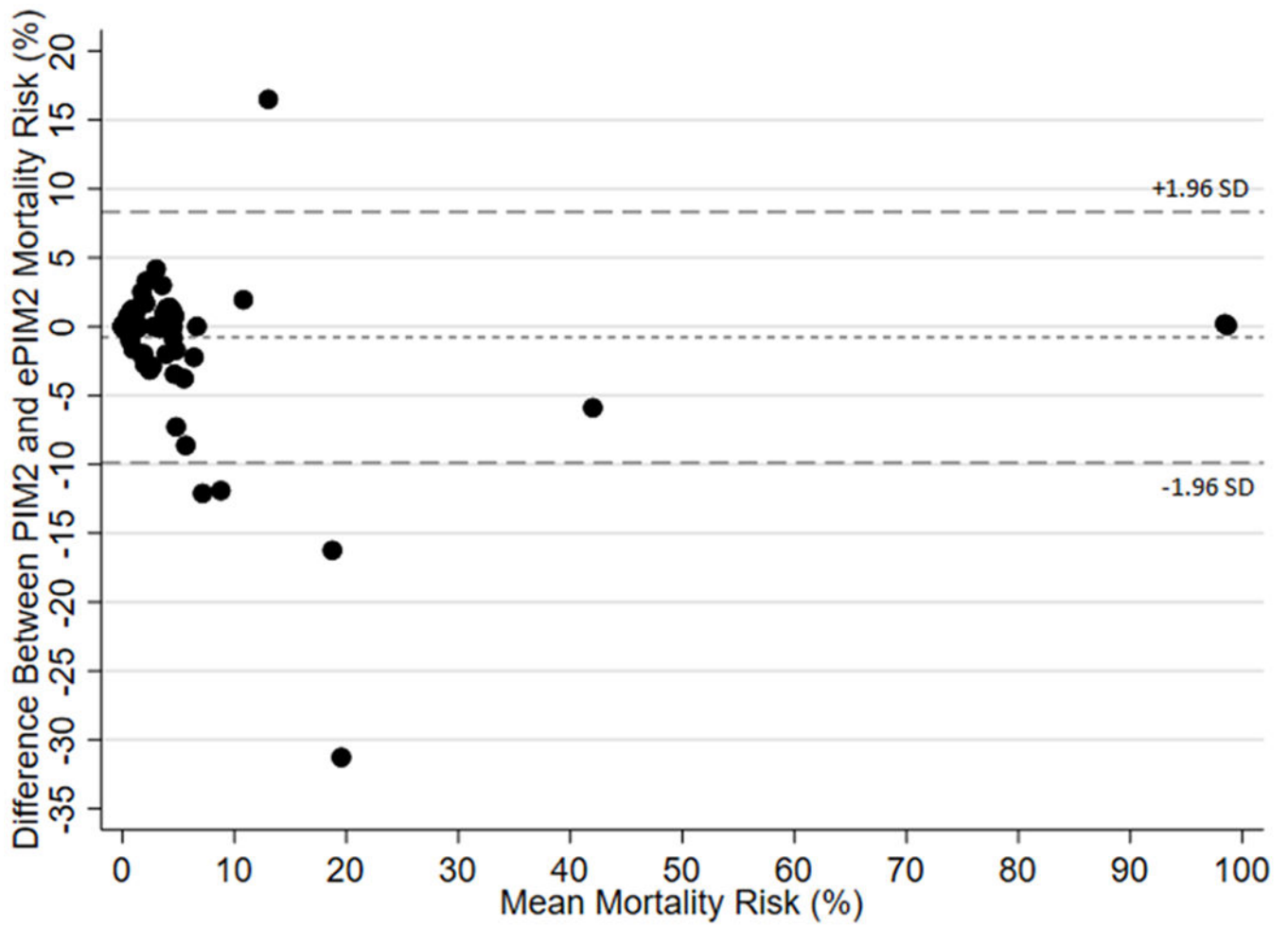


Figure 1:
Bland-Altman Plot Comparing Risk of Mortality calculated using PIM2 versus ePIM2 scores
PIM2, Pediatric Index of Mortality 2
ePIM2, Electronic Pediatric Index of Mortality 2
SD, standard deviation

Table 1:Hosmer-Lemeshow goodness-of-fit test for deciles of mortality risk: χ^2 4.19, 8 df, p=0.84

Decile of Risk	n	Observed Deaths	Expected Deaths	Observed Survivors	Expected Survivors	SMR (95% CI)	p-value
0-0.0009	1262	1	0.6	1261	1261.4	1.67 (0.083-8.22)	0.61
0.0009-0.0011	1309	2	1.3	1307	1307.7	1.54 (0.26-5.08)	0.54
0.0011-0.0017	1204	1	1.6	1203	1202.4	0.63 (0.031-3.08)	0.64
0.0017-0.0042	1292	2	3.9	1290	1288.1	0.51 (0.086-1.69)	0.34
0.0042-0.0049	1323	8	6	1315	1317	1.33 (0.62-2.53)	0.41
0.0049-0.0057	1213	7	6.4	1206	1206.6	1.09 (0.48-2.16)	0.81
0.0057-0.0091	1205	5	8.1	1200	1196.9	0.62 (0.23-1.37)	0.28
0.0091-0.0181	1287	17	19.5	1270	1267.5	0.87 (0.52-1.37)	0.57
0.0181-0.0245	1232	25	25.7	1207	1206.3	0.97 (0.64-1.41)	0.89
0.0245-0.9738	1255	138	132.7	1117	1122.3	1.04 (0.88-1.23)	0.65