OBSERVATIONAL STUDY

OPEN

Performance of Pediatric Risk of Mortality IV in Brazilian PICUs: A Multicenter Prospective Study

IMPORTANCE: This is the first Brazilian study evaluating the performance of Pediatric Risk of Mortality (PRISM) IV and the first to use the calibration belt technique.

OBJECTIVES: This study aimed to evaluate the performance of PRISM IV in a large cohort of patients admitted to Brazilian PICUs.

DESIGN, SETTING AND PARTICIPANTS: This is a longitudinal, prospective, multicenter study conducted in 36 Brazilian PICUs with children between 29 days and 18 years old admitted from March 2020 to March 2022.

MAIN OUTCOMES AND MEASURES: PRISM IV's performance was assessed using the standardized mortality ratio (SMR), the area under the receiver operating characteristic curve (AUROC) with 95% CI, and the calibration belt with 80% and 95% CI.

RESULTS: A total of 12,046 patients from 36 PICUs were included. Observed overall in-hospital mortality was higher than predicted: observed = 249 (2.1%) × predicted = 188.1 (1.56%) (SMR = 1.32 [95% CI, 1.16–1.50]); discrimination was good (AUROC = 0.86 [95% CI, 0.83–0.89]), and calibration was poor, underestimating mortality over a wide range of predicted mortality (2–61%). To explore the impact of the COVID-19 pandemic on PRISM IV's performance, we divided the study period into prevaccine and postvaccine. In the prevaccine period, the SMR was 1.38 (95% CI, 1.17–1.62), the AUROC was 0.84 (95% CI, 0.80–0.88), and the range of miscalibration was broader than in the total cohort (underestimation in the 2–98% range). In the postvaccine period, the SMR was 1.26 (95% CI, 1.03–1.51), the AUROC was 0.90 (95% CI, 0.86–0.94), and the calibration belt underestimated mortality in a narrower range of 3–46% of predicted mortality.

CONCLUSIONS AND RELEVANCE: PRISM IV showed good discrimination but miscalibration across a wide range of predicted mortality and different COVID-19 pandemic periods in a large cohort. Further research with subgroup analyses are needed to develop strategies to improve the performance of PRISM IV in different and heterogeneous Brazilian healthcare contexts.

KEYWORDS: mortality; pediatric intensive care units; Pediatric Risk of Mortality IV; prognostic score; scoring system

Gustavo Rodrigues-Santos, MD, MSc[1,2]

Arnaldo Prata-Barbosa, MD, PhD[2,3]

Fernanda Lima-Setta, MD, PhD[2,4]

Pedro Henrique Nunes Costa Silami, MD, MSc[2]

Mariana Barros Genuíno de Oliveira, PT, MSc[2]

Jaqueline Rodrigues Robaina, PSYCH, PhD[2]

José Colleti Júnior, MD, PhD[5,6]

Felipe Rezende Caino de Oliveira[7]

Luís Fernando Andrade de Carvalho, MD, MSc[8,9]

Mariana Digiovanni, MD, MSc[10]

Ana Paula Novaes Bellinat, MD[11]

Thiago Peres da Silva, MD[12]

Taisa Roberta Ramos Nantes de Castilho, MD[13]

Simone Camera Gregory, MD[14]

Ana Carolina Cabral Pinheiro Scarlato, MD[15]

Paula Marins Riveiro, MD[16]

José Oliva Proença Filho, MD[17]

Antonio José Ledo Alves da Cunha, MD, PhD[2,3]

Maria Clara de Magalhães-Barbosa, MD, PhD[2]

Claudia de Souza Lopes, MD, PhD[1]

on behalf of the Brazilian Research Network in Pediatric Intensive Care (BRnet-PIC)

DOI: 10.1097/CCE.0000000000001243

The mortality rate is a commonly used indicator of the performance and quality of ICUs. One of the most widely used measures is the standardized mortality rate (SMR), which is the ratio of observed to predicted mortality (1, 2). Prognostic systems use physiologic and laboratory parameters to predict mortality based on disease severity and other risk factors (2). Currently, three predictive tools are widely used in PICUs in their updated versions: the Pediatric Risk of Mortality (PRISM) (3), the Pediatric Index of Mortality (PIM) (4), and the Pediatric Logistic Organ Dysfunction (PELOD) (5), which require varied information to calculate the risk of death.

KEY POINTS

Question: What is the performance of Pediatric Risk of Mortality (PRISM) IV in a sample from 36 Brazilian PICUs?

Findings: In this sample of Brazilian PICUs, PRISM IV underestimated the observed mortality and showed good discrimination but poor calibration in the mortality range between 2% and 61% of predicted mortality.

Meanings: Various causes of inequalities in the studied population and organizational structures of the participating PICUs may have contributed to the miscalibration range; therefore, subgroup analysis is critical for decision-making on using PRISM IV to assess the quality of care and benchmarking in Brazilian PICUs.

Regarding PRISM IV, some studies have been conducted to evaluate its use (6–8). However, these studies generally involved few institutions and included only a small number of patients. This study aimed to assess the performance of PRISM IV in a large cohort of patients admitted to Brazilian PICUs.

METHODS

This is a longitudinal, prospective, multicenter study involving 36 Brazilian PICUs. The study was carried out from March 2020 to March 2022. It was approved by the Research Ethics Committees of each participating site, which waived informed consent, as data would be treated anonymously—please see **Table S1** (http://links.lww.com/CCX/B489) for more information.

Patients were assessed for eligibility daily by the local coordinator. The inclusion and exclusion criteria and the variables required to calculate the score were the same as those of the original PRISM IV study (3). In addition, demographic information and patient outcome data (including PICU length of stay and in-hospital mortality) were collected. Cases with missing in-hospital mortality data were excluded. A spreadsheet for calculating PRISM IV and forms specifically developed for filling out unit characterization data and patient clinical data were made available to participating centers, and data submission was monitored monthly. The characteristics of the study population were compared with the original study.

To evaluate PRISM IV performance, we estimated: 1) the standardized mortality ratio (SMR), consisting of the ratio between observed and predicted mortality rates (9, 10); 2) the discrimination capacity, assessing the area under the ROC curve (AUROC) with a 95% CI (11–13); and 3) the calibration of the score along the predicted risk of mortality range, using the calibration belt, a method proposed by the *Italian Group for the Evaluation of Interventions in Intensive Care Medicine*, which consists of a function that generates a calibration curve, resulting from the relationship between each predicted (on the x-axis) and observed (y-axis) mortality rate, with areas of 80% and 95% CIs. In this method, a deviation from the bisector was considered statistically significant when the 95% CI limits did not contain the bisector (13–16). We used the R package "givitiR" for these analyses. The R 4.0.3 program (R Foundation for Statistical Computing, Vienna, Austria) was used for all statistical analyses.

During the study period, we observed that the COVID-19 pandemic could be altering the case mix of admitted patients, the organizational structures of the participating PICUs, and the mortality rate. Consequently, the performance of PRISM IV could also be affected, especially in the first wave, before the start of the vaccination campaign in Brazil (from March 2020 to January 2021). Therefore, we also stratified the PRISM IV performance analysis in two periods—before and after the SARS-CoV-2 vaccination—to explore the effect of the pandemic on the score's performance.

This study was approved by the research ethics committee of all participating PICUs, which waived the informed consent. All procedures followed were in accordance with these committees' ethical standards and the 1975 Declaration of Helsinki, as mostly recently amended. For more information, see Table S1 (http://links.lww.com/CCX/B489).

RESULTS

Out of the 13,391 eligible patients included in the study, 1,345 were excluded for the following reasons: 1,017 because of missing data related to in-hospital mortality, 83 because they were older than 18 years

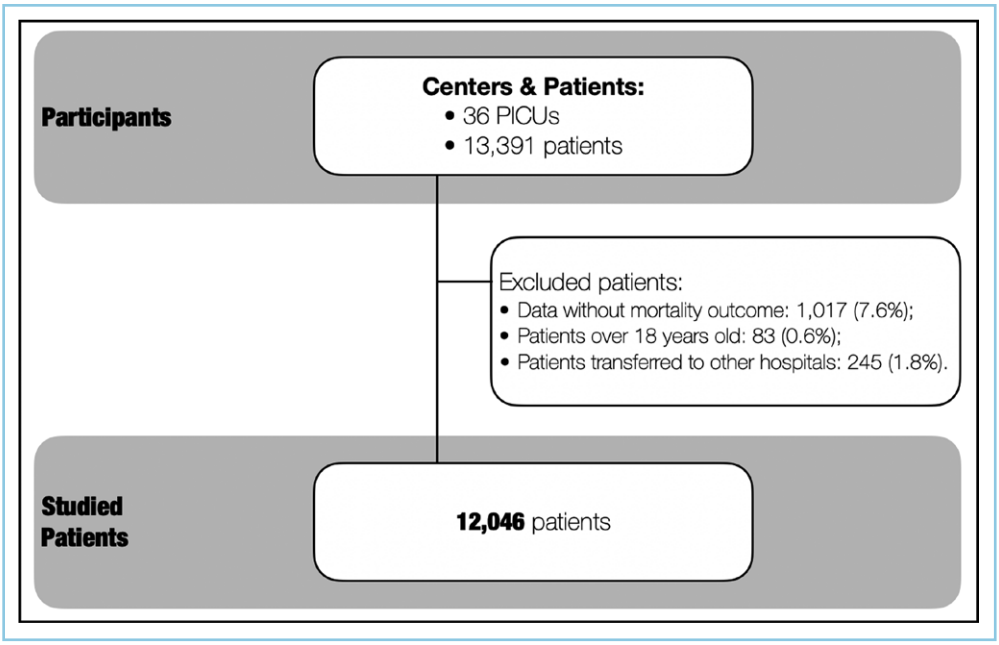

Figure 1. Flowchart of the participating PICUs and patients included in the study.

old, and 245 because they were transferred from the PICUs to other hospitals (**Fig. 1**). Therefore, the study population consisted of 12,046 patients.

Tables 1 and **2** show the characteristics of the study population and the population of the original study conducted by Pollack et al (3), which updated the PRISM IV. The median age of the study population was 2.8 years, with 41.3% under 2 years old; 54% were male subjects, 77.3% from private PICUs, and 66.2% from exclusively PICUs (33.8% were mixed units, admitting also neonates in a special area in the same unit); 89.5% were emergency (nonelective) admissions and 10.7% were surgical cases. The main source of patient admission was the emergency department (68.4%), and the three most frequent primary dysfunctions were respiratory (43.0%), neurologic (16.7%), and cardiovascular (10.6%). The median length of stay in the PICU was 3 days, the observed mortality was 249 (2.1%) patients, and the mortality predicted by PRISM IV was 188.1 (1.56%) patients (Table 2). Compared with the population of the original study, the median age of the study population was slightly lower (2.8 vs. 3.7 yr); the proportions of government-funding patients (22.7% vs. 53.8%), elective patients (10.5% vs. 36.4%), surgical patients (10.7% of surgical cases vs. 37% of postoperative patients), and cardiovascular dysfunction (10.6% vs. 24%) were also lower, as was the median length of PICU stay (3 vs. 4.9 d).

The observed in-hospital mortality was higher than predicted, with an SMR of 1.3 (95% CI, 1.2–1.5). The discrimination assessment showed an AUROC of 0.86 (95% CI, 0.83–0.89) (**Fig. 2*A***). The calibration belt underestimated mortality in the 2–61% range of mortality risk (Fig. 2*B*). Given that the predicted mortality rate in our cohort was 0.9 (Table 2), we focused on **Figure 2*B*** to examine the calibration within the predicted mortality range of 0–5% (**Fig. 2*C***). In this risk range, good calibration was observed below 2%, corresponding to 10,444 patients (86.7%), the vast majority of the study population.

Performance analysis in the pre- and post-anti-SARS-CoV-2 vaccine periods showed an SMR of 1.38 (95% CI, 1.17–1.62) and 1.26 (95% CI, 1.03–1.51), and an AUROC of 0.84 (95% CI, 0.80–0.88) and 0.90 (95% CI, 0.86–.0.94) (**Fig. 3, *A*** and ***C***), respectively. Calibration analysis indicated that the score underestimated mortality in the range of 2–98% of predicted mortality in the prevaccine period and 3–46% in the postvaccine period (**Fig. 3, *B*** and ***D***).

DISCUSSION

In this multicenter study in a large Brazilian cohort, the performance of the PRISM IV showed good discrimination but poor calibration in a wide range of predicted mortality. These results are in line with those of other studies that evaluated PRISM outside its development setting and found good discrimination but poor calibration (6, 17, 18). However, this study is unique in the pediatric intensive care setting because it used a statistical analysis technique, the calibration belt, which allowed us to demonstrate the specific range of mortality risk in which calibration was poor. Differences related to population case mix, organizational structure of PICUs, and local health policies may have contributed to the results.

Our study population exhibited similarities and differences compared with the original PRISM IV

TABLE 1.
Demographic Data and Characteristics of Patients Admitted to PICUs Compared With the Original Study (3)

Characteristics	This Study Cohort (*n* = 12,046)	Original Study, 2016[a] (*n* = 10,078)	Original Study, 2016[b] (*n* = 7,560)
Gender, *n* (%)			
Female	5,540 (46.0)	NA	4,140 (54.8)
Male	6,506 (54.0)	NA	3,420 (45.2)
Age (yr), median (interquartile range)	2.8 (0.8–7.3)	3.7 (3.2–4.1)	NA
Infants[c] (< 2 y)	4,970 (41.3)	NA	2,095 (27.7)
Preschoolers[c] (2–5 yr)	3,424 (28.4)	NA	2,142 (28.3)
Grade schoolers (6–12 yr)	2,582 (21.4)	NA	1,658 (21.9)
Adolescents (13–18 yr)	1,070 (8.9)	NA	1,665 (22.0)
Payer, *n* (%)			
Private funding (PICUs in private hospitals)	9,312 (77.3)	4,168[d] (41.4)	NA
Government funding (PICUs in public hospitals)	2,734 (22.7)	5,420[d] (53.8)	NA
PICU type, *n* (%)			
Pediatric	7,977 (66.2)	NA	NA
Mixed (pediatric and neonatal)	4,069 (33.8)	NA	NA
Admissions in PICUs with postgraduate programs, *n* (%)	6,695 (55.6)	NA	NA
Admission type, *n* (%)			
Emergency	10,786 (89.5)	6,411 (63.6)	NA
Elective	1,260 (10.5)	3,667 (36.4)	NA
Diagnostic group, *n* (%)			
Medical	10,757 (89.3)	6,281 (62.3)	NA
Surgical	1,289 (10.7)	3,797 (37.7)	NA
Cardiac surgery	257 (2.1)	1,549 (15.4)	NA

NA = not available.

[a]Total cohort of patients from the original study (Pollack et al (3)).

[b]Original study development set.

[c]In the original study, the age range was cut from 0 to 12 mo and from 1 to 5 yr.

[d]The sum does not represent 100% because of uninformed data.

validation study (3) and other evaluations of its performance (6, 18–20). The observed mortality rate was slightly lower than in the original study (2.1% vs. 2.7%) (3), and lower than in studies conducted in Turkey (8.2%) (6) and China (6.2% and 4.7%) (19, 20). The most frequent organ dysfunctions upon admission were respiratory, neurologic, and cardiovascular, consistent with other studies (6, 19, 20). However, cardiovascular dysfunction was less prevalent than in the original study (3), (2.1% vs. 15.4% in our study). Additionally, our cohort showed more respiratory dysfunctions and less elective hospitalizations than other studies (3, 6, 19), possibly because of the COVID-19 pandemic.

The SMR revealed that PRISM IV underestimated overall mortality. A study involving four PICUs in Turkey (6) showed an SMR of 1.5 (1.10–1.97), indicating a similar outcome. Conversely, a study in China described that PRISM IV provided an accurate estimation of mortality with an SMR of 1.05 (0.92–1.21) (19).

TABLE 2.
Other Demographic Data and Characteristics of Patients Admitted to PICUs Compared With the Original Study (3)

Characteristics	This Study Cohort (*n* = 12,046)	Original Study, 2016[a] (*n* = 10,078)	Original Study, 2016[b] (*n* = 7,560)
Source of PICU admission, *n* (%)			
Emergency department	8,237 (68.4)	NA	2,416 (32.0)
Ward/individual room	1,041 (8.6)	NA	828 (11.0)
Operating room	1,289 (10.7)	NA	2,834 (37.5)
Transfer from other hospital	1,474 (12.2)	NA	1,482 (19.6)
Not informed	6 (0.0)	NA	-
Primary system of dysfunction, *n* (%)			
Respiratory	5,176 (43.0)	NA	2,540 (33.6)
Neurologic	2,011 (16.7)	NA	1,508 (19.9)
Cardiovascular	1,273 (10.6)	NA	1,815 (24.0)
Hematological	510 (4.2)	NA	60 (0.8)
Endocrine	415 (3.4)	NA	250 (3.3)
Musculoskeletal	344 (2.9)	NA	319 (4.2)
Renal	330 (2.7)	NA	68 (0.9)
Not present	1,939 (16.1)	NA	-
Not informed	48 (0.4)	NA	738[c] (9.8)
Cardiac arrest before PICU admission[d], *n* (%)	122 (1.0)	141 (1.4)	NA
PICU length of stay (d), median (IQR)	3.0 (2.0–6.0)	4.9 (4.0–7.0)	NA
Total PRISM score, mean (SD)	2.35 (3.56)	NA	NA
PRISM neurologic score, mean (SD)	0.11 (1.03)	NA	NA
PRISM nonneurologic score, mean (SD)	2.24 (3.27)	NA	NA
Unadjusted mortality rate (%), median (IQR)	0.9 (0.6–1.4)	2.7 (1.3–5.0)	NA
Observed deaths (O)	249 (2.1)	275 (2.7)	NA
Expected deaths (E)	188.1	280.3	NA
SMR (O/E)[e] (95% CI)	1.3 (1.2–1.5)	1.0 (0.9–1.1)	NA

IQR = interquartile range, NA = not available, PRISM = Pediatric Risk of Mortality.

[a]Total cohort of patients from the original study (Pollack et al (3)).

[b]Original study development set.

[c]Refers to other dysfunctions.

[d]Closed chest cardiac massage occurring within 24 h before hospital admission or during hospitalization before PICU admission.

[e]Mid-*p* exact test: death vs. survival.

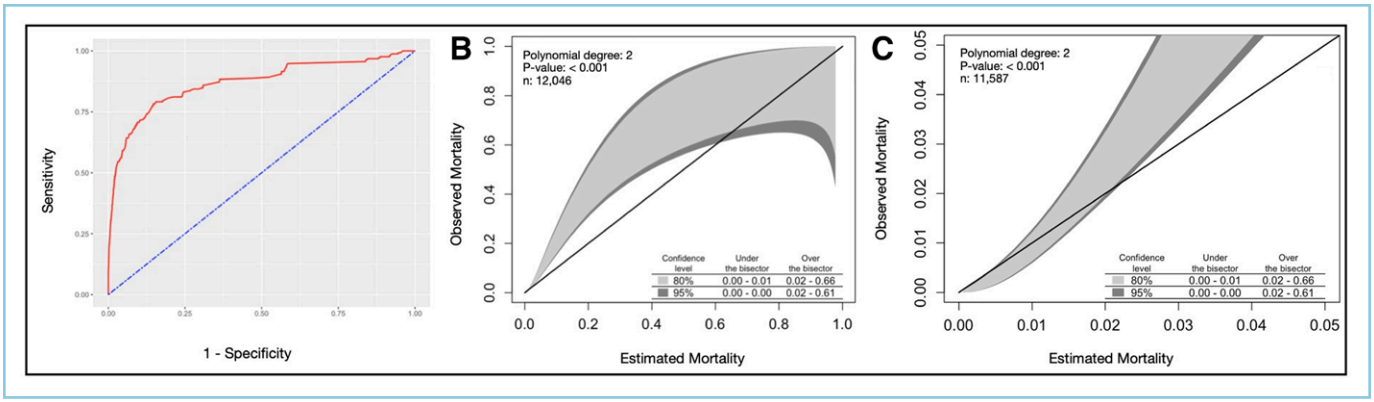

Figure 2. Pediatric Risk of Mortality IV performance evaluated by discrimination and calibration. **A**, Area under the receiver operating characteristic curve (AUROC). **B**, Calibration belt assessing the concordance of observed vs. expected mortality in the whole cohort. The inner light gray area marks the 80% CI boundary, and the dark gray outer area marks the 95% CI boundary. **C**, Amplification of the calibration belt in the risk range of predicted mortality below 5% (zooming in on part of curve **B**), which corresponds to the vast majority of the study population.

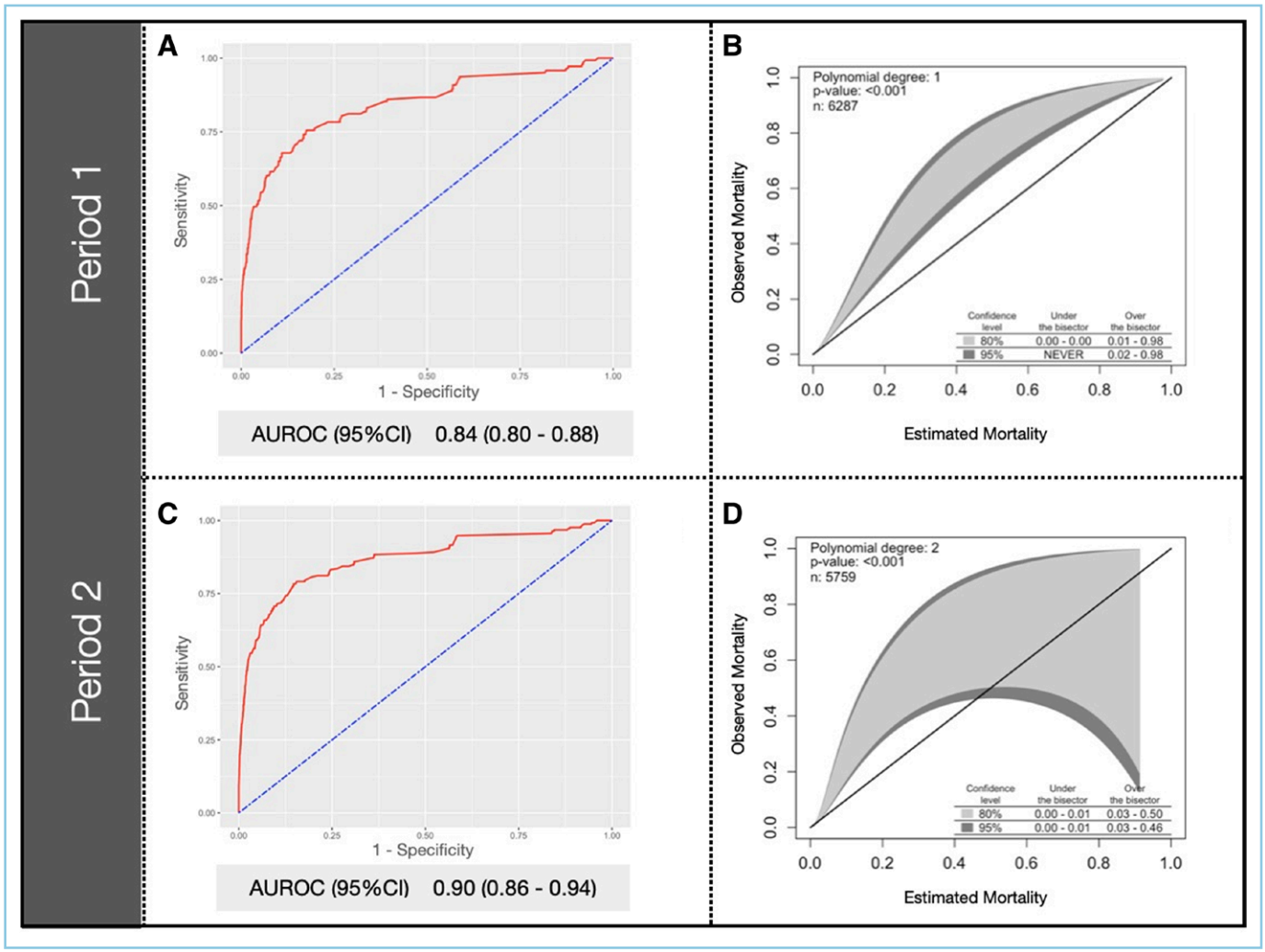

Figure 3. Pediatric Risk of Mortality IV performance evaluated by discrimination and calibration in the two COVID-19 periods: (1) prevaccination and (2) postvaccination. **A** and **C**, Area under the receiver operating characteristic curve (AUROC). **B** and **D**, Calibration belt assessing the concordance of observed vs. expected mortality in the whole cohort. The inner light gray area marks the 80% CI boundary, and the dark gray outer area marks the 95% CI boundary.

The ability to discriminate between survivors and nonsurvivors in the general population was good. This result is consistent with the conclusions of a meta-analysis published in 2021, which assessed the prediction of mortality rates of three different scores, including PRISM III/IV. This meta-analysis estimated an overall AUROC (summary operating point) of 0.84 (95% CI, 0.80–0.87) for PRISM III/IV (17). However, it is important to note that of the 18 studies on PRISM score, the only one on PRISM IV assessed the performance of an adapted version of the tool for cancer patients. A multicenter study including eight PICUs in China showed an AUROC of 0.76 (95% CI, 0.73–0.80) (20). Two other multicenter studies on the Turkish and Chinese populations showed even better discrimination estimates, with an AUROC of 0.926 (95% CI, 0.88–0.97) and 0.91 (95% CI, 0.88–0.94), respectively (6, 19).

To our knowledge, this is the first study to use the calibration belt to evaluate the performance of PRISM IV. This statistical technique is recommended for analyzing very large samples, as it avoids inconsistencies that can occur when using the Hosmer-Lemeshow method (21), and it also indicates the range of risk of death where miscalibration occurs. It has been used in a few studies evaluating the performance of mortality predictor scores in adult ICUs (22–24). Only a few pediatric studies have used this method, including two studies on PIM III (18, 25), one of which also evaluated PRISM III (18). Our results indicated that the score underestimated mortality in the range of 2–61% of predicted mortality, demonstrating poor calibration in this range of death risk. However, the score worked well in the predicted mortality range less than 2%, which encompasses most patients admitted to the participating PICUs. It also demonstrated good calibration among very critically ill patients with a risk of death above 60%. Ekinci et al (6) observed poor calibration of PRISM IV using the Hosmer-Lemeshow technique ($p = 0.008$) in Turkey, whereas Zhang et al (19) showed good calibration using the same method ($p = 0.79$) in China. The study population in China was more similar to the population of the original study by Pollack et al (3), with higher percentages of surgical patients (45.4%) and patients with cardiac dysfunction (22.3%) than the Turkish study.

The start of this study coincided with the onset of the COVID-19 pandemic in Brazil. During this period, there was an overall decrease in pediatric hospitalizations (26). Admissions of children because of COVID-19 were infrequent, around 2.5–4.1% (27). The profile of pediatric patients in the PICU was also affected, with fewer admissions, a lower proportion of respiratory causes, and shorter length of stay. However, mortality rates remained around 0.7–3% (28, 29). Nevertheless, PRISM IV still showed good discrimination in both the prevaccination and postvaccination periods. The calibration was slightly better in the postvaccination period for the risk of death more than 46%, but it still underestimated mortality in the lower risk range of 2–46%. The pandemic may have affected the case mix and the performance of the scoring system.

We must be cautious in interpreting the results of the present study. First, because the study was overlapped by the COVID-19 pandemic and this may have influenced the case mix of PICUs, especially in the prevaccination period. Second, Brazil is a continental and socioeconomically heterogeneous country. The primary health status of the population, as well as the organizational structure of the health facilities, can greatly vary. Furthermore, despite the large sample size, including PICUs from nine states and five regions of the country, it is not representative of all Brazilian PICUs. Should we use the original PRISM IV to benchmark the quality of care in Brazilian PICUs? Do these findings mean that the healthcare quality in the participant PICUs needs to improve? Do they mean PRISM IV should be customized to predict mortality in our healthcare context better? Do they mean that customization should be conducted by specific subgroups based on the inequalities of healthcare conditions of the population and healthcare facilities? Further research, including subgroup analysis, representative samples, and process-of-care analysis, is necessary to answer these questions and make decisions about using PRISM IV in Brazilian PICUs. If miscalibration stems from population differences, the score needs regional adaptation; if it reflects care quality variations, quality improvement initiatives may be warranted.

One of the strengths of this study was the use of the calibration belt. This statistical method allowed us to identify in which predicted mortality range the score was most out of calibration and demonstrate that the score had good calibration for the range where most patients admitted to the PICU are found. Although this study did not allow us to identify the reasons for

the miscalibration in this range of more severe patients, this finding points to the need for specific research with this objective. In this sense, a recently published study (30) showed that factors such as the implementation and application of managed protocols and the presence of a specialized and dedicated multidisciplinary team could contribute to reducing care-related events and mortality.

Another strength is the large number of centers and patients included. Although the sample is not representative of all Brazilian PICUs, it included private and public units from nine states and all five regions of the country. On the other hand, as a limitation, 7.6% of the patients had missing mortality data. However, we believe these losses were at random, since most of them occurred in three PICUs where there were changes in coordination with consequent non-adherence to the research protocol. As another limitation, we could not evaluate the reasons for the miscalibration observed in a wide risk range between 2% and 61% of predicted mortality. Several factors may be responsible for these findings: factors related to temporary differences in the case mix, such as the COVID-19 pandemic, or constant differences between high-income countries and low- and middle-income countries, such as the basic health conditions of the population; factors related to the healthcare system, such as limited access of the population to intensive care facilities and low availability of resources in several participant PICUs, both associated with increased mortality (31); and organizational factors related to human resources and protocols available in PICUs, also associated to patient outcomes (32). However, none of these factors could be addressed in the present study.

CONCLUSIONS

In conclusion, this first Brazilian study evaluating the performance of PRISM IV in a large sample of patients showed good discrimination but irregular calibration, with adequate adjustment for the extreme mortality risk ranges (<2% and >61%), but poor calibration between 2% and 61% mortality risk. To better understand these results and decide on the appropriate use of PRISM IV in Brazil, further research is needed to examine the behavior of the score in subgroups based on differences in case mix and in resources and organizational structures of Brazilian PICUs.

ACKNOWLEDGMENTS

We thank the following contributing authors from the Brazilian Research Network in Pediatric Intensive Care (BRnet-PIC): Lenira Medeiros de Morais Daibes Rachid (Hospital Copa D`Or, Rio de Janeiro, RJ, Brazil); Igor Bromonschenkel Brandão (Hospital Niteroi D`Or, Rio de Janeiro, RJ, Brazil); Lucas Pulcheri (Hospital Niteroi D`Or, Rio de Janeiro, RJ, Brazil); Thaís de Mello Cesar Bernardi (Hospital da Criança de Jabaquara); José Carlos Fernandes (Associação Beneficiente Síria–Hospital do Coração (Hcor), São Paulo, SP, Brazil); Luiz Aurelio de Oliveira (Hospital Ribeirão Pires, Ribeirão Pires, SP, Brazil); Evane Cordeiro Franco (Hospital e Maternidade Brasil, Santo André, SP, Brazil); Cibele Cristina Manzoni Ribeiro Borsett (Hospital São Luiz São Caetano; São Caetano, BA, Brazil); Roberto Sapolnik (Hospital São Rafael, Salvador, BA, Brazil); Diogo Pedroso (Hospital Santa Luzia, Brasília, DF, Brazil); Thallys Ramalho Suzart Alves (Hospital Santa Helena, Brasília, DF, Brazil); Alessandra Geisler Daud Lopes (Hospital Municipal Infantil Menino Jesus, São Paulo, SP, Brazil); Letícia Alves Andrade Albuquerque (Hospital Infantil João Paulo II, Belo Horizonte, MG, Brazil); Lara de Araújo Torreão (Hospital Aliança, Salvador, BA, Brazil); Márcio Miranda Brito (Hospital Municipal de Araguaína, Araguaína, TO, Brazil); Carolina Friedrich Amoretti (Complexo Hospitalar Universitário Prof. Edgard Santos, Salvador, BA, Brazil); Graziela de Araújo Costa (Hospital Sírio Libanês, São Paulo, SP, Brazil); Jurema Amancio Mascarenhas (Hospital Martagão Gesteira, Salvador, BA, Brazil); Wendell Paiva Vita (Hospital Universitario Evangélico Mackenzie, Curitiba, PR, Brazil).

1 *Department of Epidemiology, Institute of Social Medicine, State University of Rio de Janeiro, Brazil.*

2 *Department of Pediatrics, D'Or Institute for Research and Education (IDOR), Rio de Janeiro, RJ, Brazil.*

3 *Martagão Gesteira Institute of Pediatrics and Child Care, Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil.*

4 *Pediatric Intensive Care Unit, Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira, FIOCRUZ, Rio de Janeiro, RJ, Brazil.*

5 *Pediatric Intensive Care Unit, Hospital Assunção, São Bernardo do Campo, SP, Brazil.*

6 *Department of Pediatrics, Faculdade de Medicina de Jundiaí, Jundiaí, SP, Brazil.*

7 *Pediatric Intensive Care Unit, Hospital Alvorada Moema, Moema, SP, Brazil.*

8 *Pediatric Intensive Care Unit, Hospital João Paulo II, Belo Horizonte, MG, Brazil.*

9 *Pediatric Intensive Care Unit, Hospital João XXIII Hospital, Belo Horizonte, MG, Brazil.*

10 *Pediatric Intensive Care Unit, Hospital Universitario Evangélico Mackenzie, Curitiba, PR, Brazil.*

11 *Pediatric Intensive Care Unit, Hospital Martagão Gesteira, Salvador, BA, Brazil.*

12 *Pediatric Intensive Care Unit, Hospital Oeste D`Or, Rio de Janeiro, RJ, Brazil.*

13 *Pediatric Intensive Care Unit, Hospital Anália Franco, São Paulo, SP, Brazil.*

14 *Pediatric Intensive Care Unit, Hospital Estadual da Criança, Rio de Janeiro, RJ, Brazil.*

15 *Pediatric Intensive Care Unit, Hospital Rios D`Or, Rio de Janeiro, RJ, Brazil.*

16 *Pediatric Intensive Care Unit, Hospital Caxias D`Or, Rio de Janeiro, RJ, Brazil.*

17 *Pediatric Intensive Care Unit, Hospital e Maternidade Brasil, Santo André, SP, Brazil.*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccejournal).

Supported, in part, by funds from the Department of Pediatrics of D'Or Institute for Research and Education.

The authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: gustavo.rodrigues@idor.org

Please see the Brazilian Research Network in Pediatric Intensive Care (BRnet-PIC) in the Acknowledgments.

This work was performed at the Department of Epidemiology, Institute of Social Medicine, State University of Rio de Janeir, Rio de Janeiro, RJ, Brazil.

REFERENCES

1. Zimmerman JE, Kramer AA: A history of outcome prediction in the ICU. *Curr Opin Crit Care* 2014; 20:550–556
2. Kramer AA, Higgins TL, Zimmerman JE: Comparing observed and predicted mortality among ICUs using different prognostic systems: Why do performance assessments differ? *Crit Care Med* 2015; 43:261–269
3. Pollack MM, Holubkov R, Funai T, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Collaborative Pediatric Critical Care Research Network: The pediatric risk of mortality score. *Pediatr Crit Care Med* 2016; 17:2–9
4. Straney L, Clements A, Parslow RC, et al; ANZICS Paediatric Study Group and the Paediatric Intensive Care Audit Network: Paediatric index of mortality 3: An updated model for predicting mortality in pediatric intensive care. *Pediatr Crit Care Med* 2013; 14:673–681
5. Leteurtre S, Duhamel A, Salleron J, et al; Groupe Francophone de Réanimation et d'Urgences Pédiatriques (GFRUP): PELOD-2: An update of the pediatric logistic organ dysfunction score. *Crit Care Med* 2013; 41:1761–1773
6. Ekinci F, Yildizdas D, Horoz OO, et al: Performance and analysis of four pediatric mortality prediction scores among critically ill children: A multicenter prospective observational study in four PICUs. *Arch Pediatr* 2022; 29:407–414
7. Niederwanger C, Varga T, Hell T, et al: Comparison of pediatric scoring systems for mortality in septic patients and the impact of missing information on their predictive power: A retrospective analysis. *PeerJ* 2020; 8:e9993
8. Leal PDB, De Araujo OR, Petrilli AS, et al: PRISM 4-C: An adapted PRISM IV algorithm for children with cancer. *J Pediatr Hematol Oncol* 2020; 42:e563–e568
9. Vandenbroucke JP: A shortcut method for calculating the 95 per cent confidence interval of the standardized mortality ratio. *Am J Epidemiol* 1982; 115:303–304
10. Naing NN: Easy way to learn to standardization: Direct and indirect methods. *Malays J Med Sci* 2000; 7:10–15
11. Bewick V, Cheek L, Ball J: Statistics review 13: Receiver operating characteristics curves. *Crit Care* 2004; 8:508–512
12. Murphy-Filkins RL, Teres D, Lemeshow S, et al: Effect of changing patient mix on the performance of an intensive care unit severity-of-illness model: How to distinguish a general from a specialty intensive care unit. *Crit Care Med* 1996; 24:1968–1973
13. Alba AC, Agoritsas T, Walsh M, et al: Discrimination and calibration of clinical prediction models: Users' guides to the medical literature. *JAMA* 2017; 318:1377–1384
14. Nattino G, Finazzi S, Bertolini G: A new calibration test and a reappraisal of the calibration belt for the assessment of prediction models based on dichotomous outcomes. *Stat Med* 2014; 33:2390–2407
15. Nattino G, Finazzi S, Bertolini G: A new test and graphical tool to assess the goodness of fit of logistic regression models. *Stat Med* 2016; 35:709–720
16. Finazzi S, Poole D, Luciani D, et al: Calibration belt for quality-of-care assessment based on dichotomous outcomes. *PLoS One* 2011; 6:e16110
17. Shen Y, Jiang J: Meta-analysis for the prediction of mortality rates in a pediatric intensive care unit using different scores: PRISM-III/IV, PIM-3, and PELOD-2. *Front Pediatr* 2021; 9:712276
18. Jacobs A, Flechet M, Vanhorebeek I, et al: Performance of pediatric mortality prediction scores for PICU mortality and 90-day mortality. *Pediatr Crit Care Med* 2019; 20:113–119
19. Zhang XP, Feng YX, Li Y, et al: Performance of the PRISM I, PIM2, PELOD-2 and PRISM IV scoring systems in western China: A multicenter prospective study. *World J Pediatr* 2022; 18:818–824
20. Zhang Z, Huang X, Wang Y, et al: Performance of three mortality prediction scores and evaluation of important determinants in eight pediatric intensive care units in China. *Front Pediatr* 2020; 8:552

21. Van Calster B, McLernon DJ, van Smeden M, et al; Topic Group "Evaluating Diagnostic Tests and Prediction Models" of the STRATOS Initiative: Calibration: The Achilles heel of predictive analytics. *BMC Med* 2019; 17:230
22. Jahn M, Rekowski J, Gerken G, et al: The predictive performance of SAPS 2 and SAPS 3 in an intermediate care unit for internal medicine at a German university transplant center; A retrospective analysis. *PLoS One* 2019; 14:e0222164
23. Moralez GM, Rabello LSCF, Lisboa TC, et al; ORCHESTRA Study Investigators: External validation of SAPS 3 and MPM0-III scores in 48,816 patients from 72 Brazilian ICUs. *Ann Intensive Care* 2017; 7:53
24. Taniguchi LU, Siqueira EMP: Comparison of SAPS 3 performance in patients with and without solid tumor admitted to an intensive care unit in Brazil: A retrospective cohort study. *Rev Bras Ter Intensiva* 2020; 32:521–527
25. Genu DHS, Lima-Setta F, Colleti J, et al; Brazilian Research Network in Pediatric Intensive Care (BRnet-PIC): Multicenter validation of PIM3 and PIM2 in Brazilian pediatric intensive care units. *Front Pediatr* 2022; 10:1036007
26. Foster AA, Walls TA, Alade KH, et al; ACEP Pediatric Emergency Medicine Committee: Review of pediatric emergency care and the COVID-19 pandemic. *J Am Coll Emerg Physicians Open* 2023; 4:e13073
27. Oliveira EA, Colosimo EA, Simões e Silva AC, et al: Clinical characteristics and risk factors for death among hospitalised children and adolescents with COVID-19 in Brazil: An analysis of a nationwide database. *Lancet Child Adolesc Health* 2021; 5:559–568
28. Prata-Barbosa A, Lima-Setta F, Santos GR, et al: Pediatric patients with COVID-19 admitted to intensive care units in Brazil: A prospective multicenter study. *J Pediatr (Rio J)* 2020; 96:582–592
29. Araujo OR de, Almeida CG de, Lima-Setta F, et al: The impact of the novel coronavirus on Brazilian PICUs. *Pediatr Crit Care Med* 2020; 21:1059–1063
30. Soares M, Salluh JIF, Zampieri FG, et al: A decade of the ORCHESTRA study: Organizational characteristics, patient outcomes, performance and efficiency in critical care. *Crit Care Sci* 2024; 36:e20240118en
31. Machado FR, Cavalcanti AB, Bozza FA, et al; SPREAD Investigators: The epidemiology of sepsis in Brazilian intensive care units (the Sepsis Prevalence Assessment Database, SPREAD): An observational study. *Lancet Infect Dis* 2017; 17:1180–1189
32. Soares M, Bozza FA, Angus DC, et al: Organizational characteristics, outcomes, and resource use in 78 Brazilian intensive care units: The ORCHESTRA study. *Intensive Care Med* 2015; 41:2149–2160