



Performance of the PRISM I, PIM2, PELOD-2 and PRISM IV scoring systems in western China: a multicenter prospective study

Xue-Peng Zhang^{1,2,8} · Yun-Xia Feng³ · Yang Li² · Guo-Yan Lu⁴ · Xin-Yue Zhou⁵ · Can-Zheng Wei² · Xi-Ying Gui⁶ · Kai-Ying Yang¹ · Tong Qiu¹ · Jiang-Yuan Zhou¹ · Hua Yao² · Geng Zhang² · Wen-Qi Zhang⁷ · Yu-Hang Hu⁵ · Hong Wu⁶ · Si-Yuan Chen² · Yi Ji¹

Received: 18 January 2022 / Accepted: 25 July 2022 / Published online: 13 September 2022
© Children's Hospital, Zhejiang University School of Medicine 2022

Abstract

Background The aim of this study was to evaluate the performance of the four scoring tools in predicting mortality in pediatric intensive care units (PICUs) in western China.

Methods This was a multicenter, prospective, cohort study conducted in six PICUs in western China. The performances of the scoring systems were evaluated based on both discrimination and calibration. Discrimination was assessed by calculating the area under the receiver operating characteristic curve (AUC) for each model. Calibration was measured across defined groups based on mortality risk using the Hosmer-Lemeshow goodness-of-fit test.

Results A total of 2034 patients were included in this study, of whom 127 (6.2%) died. For the entire cohort, AUCs for Pediatric Risk of Mortality Score (PRISM) I, Pediatric Index of Mortality 2 (PIM2), Pediatric Logistic Organ Dysfunction Score-2 (PELOD-2) and PRISM IV were 0.88 [95% confidence interval (CI) 0.85–0.92], 0.84 (95% CI 0.80–0.88), 0.80 (95% CI 0.75–0.85), and 0.91 (95% CI 0.88–0.94), respectively. The Hosmer-Lemeshow goodness-of-fit Chi-square value was 12.71 ($P=0.12$) for PRISM I, 4.70 ($P=0.79$) for PIM2, 205.98 ($P<0.001$) for PELOD-2, and 7.50 ($P=0.48$) for PRISM IV [degree of freedom (df)=8]. The standardized mortality ratios obtained with the PRISM I, PIM2, PELOD-2, and PRISM IV models were 0.87 (95% CI, 0.75–1.01), 0.97 (95% CI, 0.85–1.12), 1.74 (95% CI, 1.58–1.92), and 1.05 (95% CI, 0.92–1.21), respectively.

Conclusions PRISM IV performed best and can be used as a prediction tool in PICUs in Western China. However, PRISM IV needs to be further validated in NICUs.

Keywords Illness severity score · Intensive care · Mortality prediction · Pediatric

Introduction

Having tools for predicting mortality in intensive care units (ICUs) is crucial. These tools are especially necessary when assessing quality of care and comparing

✉ Si-Yuan Chen
siy_chen@163.com

✉ Yi Ji
jijiyuanyuan@163.com

¹ Department of Pediatric Surgery, West China Hospital of Sichuan University, Chengdu 610041, China

² Department of Critical Care Medicine, West China Hospital of Sichuan University, Chengdu 610041, China

³ Department of Nephrology, Mianyang Central Hospital, University of Electronic Science and Technology of China, Mianyang 621000, China

⁴ Department of Pediatric Critical Care Medicine, West China Women's and Children's Hospital, Sichuan University, Chengdu 610041, China

⁵ Department of Pediatric Critical Care Medicine, Sichuan Provincial Maternity and Child Health Care Hospital, Chengdu 610041, China

⁶ Department of Critical Care Medicine, People's Hospital of Tibet Autonomous Region, Lhasa 850000, China

⁷ Department of Anesthesiology, the Affiliated Hospital of Guizhou Medical University, Guiyang 550001, China

⁸ Department of Critical Care Medicine, Mianyang Central Hospital, University of Electronic Science and Technology of China, Mianyang 621000, China

performance among ICUs because variation in cases among ICUs is inevitable. In addition, these tools serve as clinical research tools to evaluate the severity of illness among study populations. Currently, there are three risk-adjustment tools in pediatric intensive care units (PICUs) that are widely used: the Pediatric Risk of Mortality (PRISM), Pediatric Index of Mortality (PIM) and Pediatric Logistic Organ Dysfunction (PELOD) scoring systems [1–3].

The three systems differ in the information required to calculate the risk of death. The PRISM was calculated using variables with the most abnormal values within 24 hours after admission, whereas the PIM was calculated by using the information collected on admission. The PELOD was created for assessing organ dysfunction in critically ill children, and scoring is performed by using the most abnormal values of each variable during the entire intensive care unit (ICU) stay [3]. Several studies have reported that the PELOD-2 scores on day 1 (d1PELOD-2) are strongly associated with PICU mortality [4, 5].

Before these scoring systems are applied in other populations, it is important to assess their validity in those populations. The three systems have been validated in many countries worldwide [6–11]. However, to our knowledge, there are few prospective multicenter studies that have validated these tools in western China. The current study was designed to determine the performance of the PRISM I, PIM2, PELOD-2, and PRISM IV scoring systems in a population of pediatric patients admitted to six PICUs in western China.

Methods

Setting and patients

This study was a prospective, multicenter, cohort study conducted in PICUs in six tertiary hospitals in western China between February 2018 and January 2019. All patients consecutively admitted to the PICUs were enrolled. Exclusion criteria included the following: age > 16 years; preterm birth (< 37 weeks); transferred to another PICU, and death within two hours following ICU admission. Patients readmitted to the PICU were included as new admissions only if the admission occurred more than 48 hours after transfer to another hospital ward; otherwise, the initial admissions and readmission were considered as a single admission. Patients in the PICUs at the end of the study were considered alive. Patient data were collected anonymously for privacy considerations. Each child was identified by their admission number. This study

was reviewed and approved by the Ethics Committee of the Central Processing Center (West China Hospital of Sichuan University).

Data collection

The variables with the most abnormal values within 24 hours after admission were collected for the PRISM and d1PELOD-2 score calculations. The PIM-2 scores were calculated using the information collected within one hour after admission [12]. Laboratory data were collected from two hours before PICU admission to four hours after admission, and physiological variables measured during the first four hours in the PICU were collected for PRISM IV scores [13]. Patient information, including patient age and sex, the main reason for PICU admission, the diagnosis at the time of PICU admission, length of PICU stay, and PICU mortality, also was collected. The variables with no test results were considered normal. All inconsistencies were discussed and resolved through telemeetings with one author at a central processing center. A research assistant at each study site assessed the data for accuracy, and an investigator at the central processing center monitored consistency of the data throughout the study.

Statistical analysis

Statistical analyses were performed with SPSS Statistical Package, version 22.0 (SPSS Inc., Chicago, Illinois) and R software version 3.61. Descriptive statistics were presented as the mean \pm standard deviation (SD) and as medians (interquartile range, IQR) for normally and nonnormally distributed data, respectively. The performance of each tool was evaluated by assessing the discrimination and calibration. Discrimination is the ability of a scoring system to distinguish between survivors and non-survivors correctly and was assessed using the area under the receiver operating characteristic (ROC) curve along with 95% confidence intervals (CIs). An area under curve (AUC) > 0.70 was considered indicative of acceptable discriminatory performance, AUC > 0.80 as good discrimination, and AUC > 0.90 as excellent discrimination [14]. Calibration is the ability of a scoring system to match the observed number of deaths and was assessed by using the Hosmer–Lemeshow goodness-of-fit test [15]. Ten intervals were categorized according to the predicted probability of mortality as described in previous studies, and the Chi-square statistic was calculated as $\sum (O - E)^2/E$, where O is the observed number of events and E is the expected number of events in each interval. A *P*-value > 0.05 was considered to indicate good calibration. The ratios of the observed number of deaths to the predicted number of deaths [standardized mortality ratios (SMRs)] also were calculated along with their 95% CIs. If the upper

95% CI of the SMR was < 1.0 , the observed mortality was regarded as being lower than the predicted mortality.

Results

All six hospitals were large, tertiary, referral centers located in western China and provided medical care for a population of 150 million people. Details of the six PICUs are listed in Supplementary Table 1. Two units were in children's hospitals. Half of the units had fewer than ten beds. All of the PICUs treated medical and general surgical patients.

During the study period, a total of 2282 patients were consecutively admitted to the six PICUs. Among them, 248 patients were excluded: 93 older than 16 years, 62 premature infants, 56 with incomplete data, and 37 who were discharged against medical advice. Thus, 2034 patients were enrolled in the study. The demographic and clinical characteristics of the study population are presented in Table 1. The median (IQR) age was 14 (3–51) months. Overall, the PICU mortality was 6.2%. The length of PICU stay was 4.8 (2.0–9.5) days. The largest percentage of patients (27.8%)

were grouped into the respiratory disease category. The second largest category of patients was digestive disease (24.1%), followed by cardiac disease (22.3%), neurological disease (12.8%), and injury and poisoning (2.2%).

The ROC curves that demonstrated the discrimination abilities of the systems for the entire cohort are presented in Fig. 1. The PRISM IV had the highest AUC, and all tools showed good discrimination between survival and nonsurvival. The AUCs for PRISM I, PIM2, PELOD-2, and PRISM IV were 0.88 (95% CI 0.85–0.92), 0.84 (95% CI 0.80–0.88), 0.80 (95% CI 0.75–0.85), and 0.91 (95% CI 0.88–0.94), respectively (Table 2). PRISM I, PIM2, PELOD-2, and PRISM IV predicted 145.61 (7.16%), 130.27 (6.40%), 72.90 (3.58%), and 120.50 (5.92%) deaths, respectively. The SMRs for PRISM I, PIM2, PELOD-2, and PRISM IV were 0.87 (95% CI 0.75–1.01), 0.97 (95% CI 0.85–1.12), 1.74 (95% CI 1.58–1.92), and 1.05 (95% CI 0.92–1.21), respectively (Table 2). The results of the Hosmer-Lemeshow goodness-of-fit tests with eight degrees of freedom showed that PIM2 and PRISM IV achieved the best calibration ($P = 0.79$ and $P = 0.48$), PRISM achieved good calibration ($P = 0.12$), and PELOD-2 showed a lack of fit and therefore had poor calibration ($P < 0.001$). Detailed information on the calibration of each tool across various levels of probability of death is shown in Supplementary Table 2.

In the subgroup analyses the PRISM IV and PIM2 systems achieved good calibration across all strata except for babies < 1 month of age (Fig. 2b and d). The PRISM IV overestimated mortality in babies < 1 month (14.00 vs. 26.09, SMR = 0.54), while PIM2 underestimated mortality

Table 1 Demographic and clinical characteristics of the 2034 children

Characteristics	Value
Number of patients, <i>n</i>	2034
Age (mon)	14 (3, 51)
Gender	
Male, <i>n</i> (%)	1145 (56.3%)
Female, <i>n</i> (%)	889 (43.7%)
Diagnostic group	
Medical, <i>n</i> (%)	1110 (54.6%)
Surgical, <i>n</i> (%)	924 (45.4%)
Diagnostic group	
Cardiac disease, <i>n</i> (%)	454 (22.3%)
Respiratory disease, <i>n</i> (%)	566 (27.8%)
Neurological disease, <i>n</i> (%)	261 (12.8%)
Digestive disease, <i>n</i> (%)	491 (24.1%)
Injury and poisoning, <i>n</i> (%)	44 (2.2%)
Other, <i>n</i> (%)	218 (10.7%)
Length of ICU stay (d)	4.8 (2.0, 9.5)
Ventilation, <i>n</i> (%)	1088 (53.5%)
PICU mortality, <i>n</i> (%)	127 (6.2%)
Predicted mortality	
PRISM I	7.16% (145.61)
PIM2	6.40% (130.27)
PELOD-2	3.58% (72.90)
PIRMS IV	5.92% (120.50)

ICU intensive care unit, PICU pediatric intensive care unit, PRISM pediatric risk of mortality score, PIM2 pediatric index of mortality 2, PELOD-2 pediatric logistic organ dysfunction-2

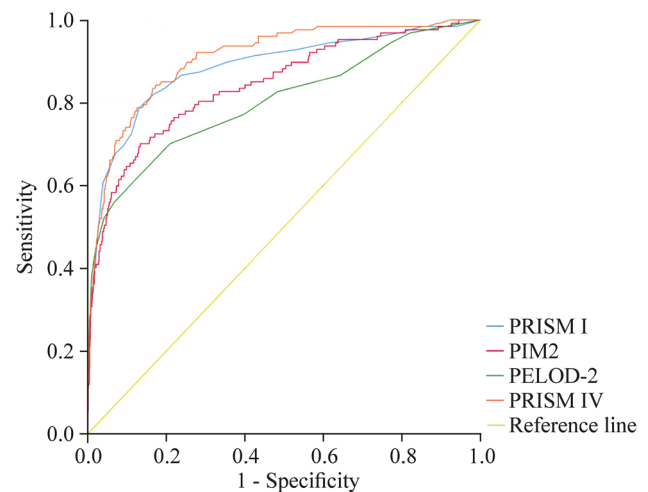


Fig. 1 ROC curves of the four scoring systems. The AUC for PRISM I was 0.88 (95% CI 0.86–0.93), for PIM2 was 0.84 (95% CI 0.80–0.88), for PELOD2 was 0.80 (95% CI 0.75–0.85), and for PRISM IV was 0.91 (95% CI 0.88–0.94). AUC area under curve, ROC receiver operating characteristic, PRISM Pediatric Risk of Mortality Score, PIM2 Pediatric Index of Mortality 2, PELOD-2 Pediatric Logistic Organ Dysfunction-2

Table 2 Model fit and discrimination in overall patients

Variables	N	Observed deaths	Expected deaths	SMR (CI 95%)	χ^2	P*	AUC (CI 95%)
PRISM I	2034	127	145.61	0.87 (0.75–1.01)	12.71	0.12	0.88 (0.85–0.92)
PIM2	2034	127	130.27	0.97 (0.85–1.12)	4.70	0.79	0.84 (0.80–0.88)
PELOD-2	2034	127	72.90	1.74 (1.58–1.92)	205.98	< 0.001	0.80 (0.75–0.85)
PRISM IV	2034	127	120.50	1.05 (0.92–1.21)	7.50	0.48	0.91 (0.88–0.94)

AUC area under receiver operating characteristic curve, SMR standardized mortality ratio, PRISM pediatric risk of mortality score, PIM2 pediatric index of mortality 2, PELOD-2 pediatric logistic organ dysfunction-2, χ^2 Chi-square value, CI confidence interval

*degrees of freedom were eight

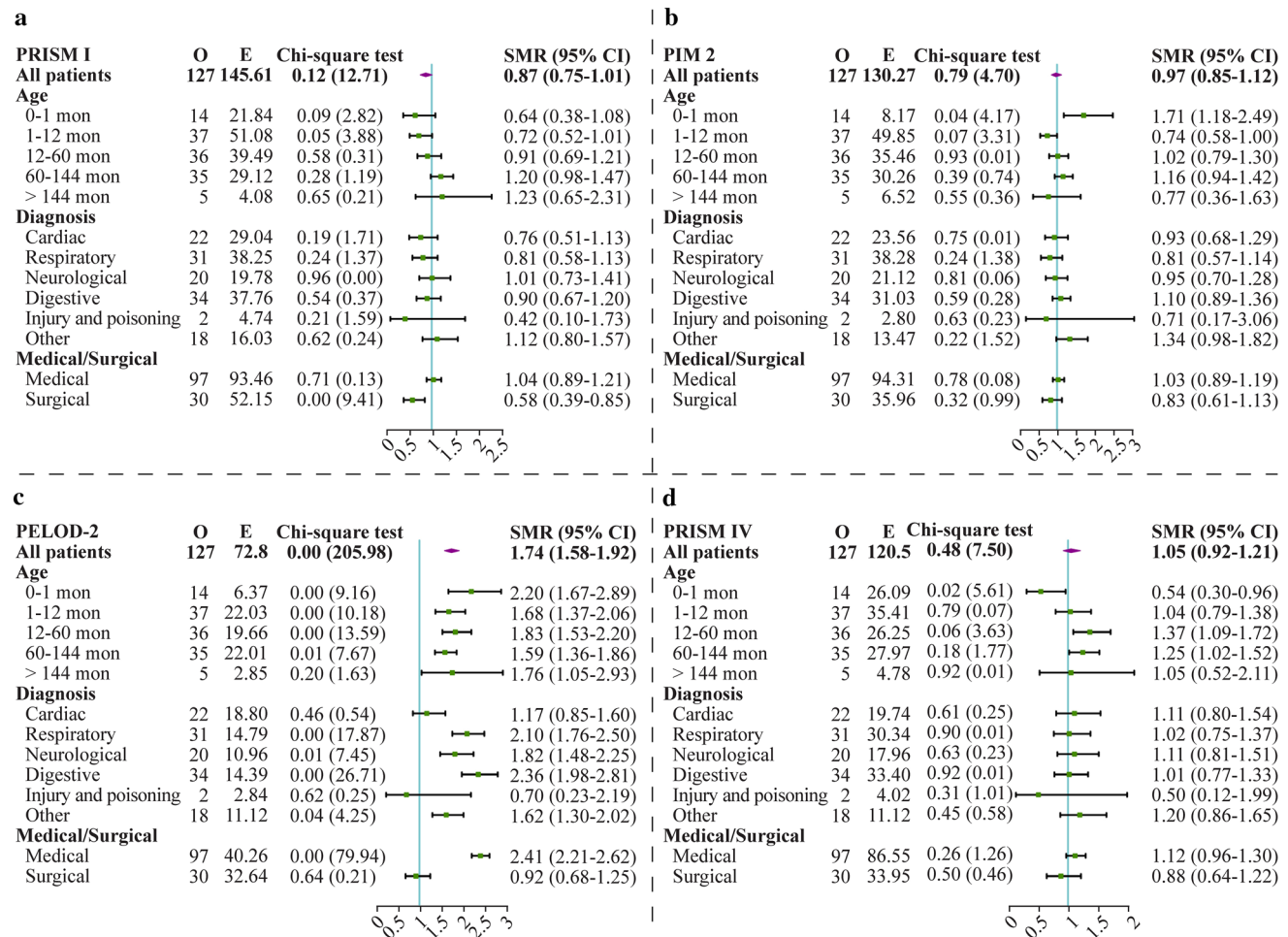


Fig. 2 Model fit by subgroups and forest plots for SMR. The results of the Chi-square test are presented as the P value (χ^2 value). The degree of freedom for the overall analysis was eight, and for each subgroup analysis, it was one. E expected deaths, O

observed deaths, SMR standardized mortality ratio, CI confidence interval, PRISM pediatric risk of mortality score, PIM2 pediatric index of mortality 2, PELOD-2 pediatric logistic organ dysfunction-2

(14.00 vs. 8.17, SMR = 1.71). PRISM I also showed good calibration except for surgical patients (Fig. 2a). The PELOD-2 score nearly underestimated mortality in all subgroups (Fig. 2c). Figure 3 shows the discrimination power of the four tools in each subgroup. PRISM I and

PRISM IV discriminated survival from nonsurvival well across all subgroups. Except for adolescent patients, PIM2 and PELOD-2 scores also could discriminate survival. In the subgroup analysis for each hospital, all tools except PELOD-2 had good discrimination, and PRISM IV had the

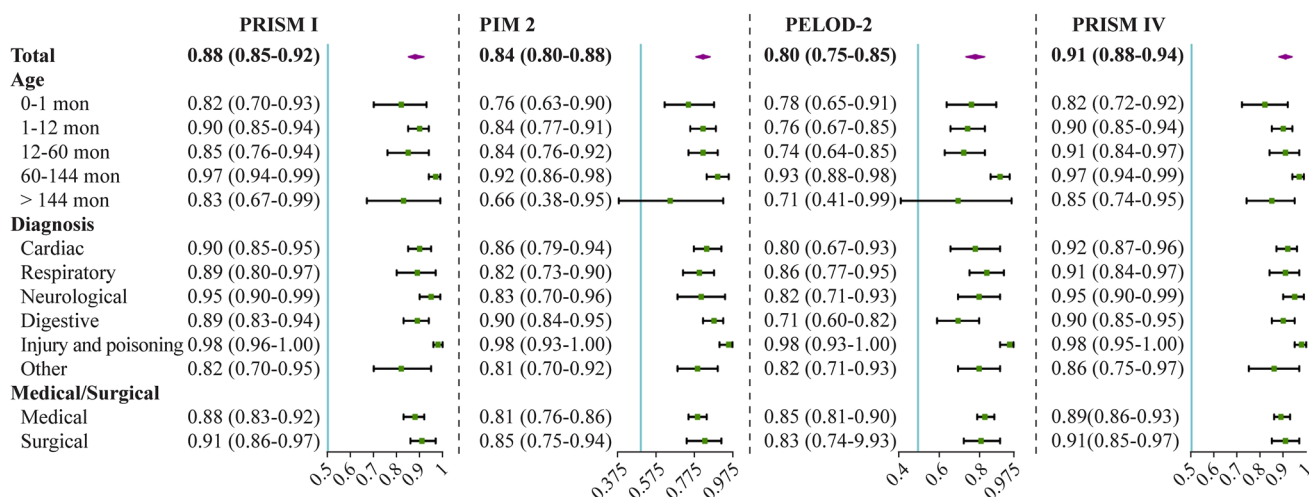


Fig. 3 Forest plots for AUCs of the four tools across each subgroup. The reference line was 0.5. *AUC* area under curve, *PRISM* pediatric risk of mortality score, *PIM2* pediatric index of mortality 2, *PELOD-2* pediatric logistic organ dysfunction-2

best calibration ($P > 0.1$ in five hospitals) (Supplementary Table 3).

Discussion

In this study, we evaluated the performance of the four scoring systems in our PICUs. The results showed that PRISM IV achieved the best discrimination and calibration among the four tools. The PIM2 and PRISM I also showed good discrimination and calibration. The PELOD-2 score had good discrimination but poor calibration.

The first three versions of PRISM need data collected 24 hours after admission to reflect illness severity. Generally, early treatment could improve the abnormal values during the first day in the ICU. Once the first measurement of value calculating PRISM occurs after early treatment, it may lead to bias [16]. When comparing the performance of two ICUs, the bias may result in the risk that the better ICU appears to perform worse because the patients with a higher risk of mortality may have lower PRISM scores in the better ICU. Therefore, one of the most important changes of the most recent version PRISM IV was the time period for calculating PRISM [13]. In addition, the algorithm used to calculate mortality risk is publicly available.

Previously, PRISM I achieved good discrimination and calibration in two cohort studies in India [6, 17]. However, PRISM underpredicted mortality in one study, possibly due to the high severity of illness and limited ICU resources in that study [6]. In China, especially western China, pediatric critical care medicine also has the problem of limited resources, which is due to medical disparities [18, 19]. Unlike the results observed in the Indian studies, the PRISM I score overpredicted mortality in our study (SMR = 0.87,

95% CI 0.75–1.01). The illness severity of patients was lower in our study than in the studies in India [7 (4, 11) vs. 16 (15, 17.4)], which may partly explain the overprediction [6]. Another possible reason is that PRISM I was derived from data collected 40 years ago. Advances in the quality of care in PICUs also could have affected the performance of the models in mortality prediction.

The most recent version of the PRISM scoring system was PRISM IV, which has not been validated in western China. Compared to PRISM I, PRISM IV showed better performance in mortality prediction in this validation study. Interestingly, the performance of PRISM IV in a recent study in eastern China was not as good as that of PRISM IV in our study [0.76 (95% CI 0.73–0.80) vs. 0.91 (95% CI 0.88–0.94)] [20]. In our study, PRISM IV overestimated mortality in babies < 1 month. The poor calibration may be explained by the following reasons. First, none of the six PICUs was a neonatal ICU (NICU). Most neonates with a high risk of mortality may be admitted to NICUs, which could lead to bias. Another possible explanation may be the underdeveloped level in western China. Patients with limited education may not be aware of the clues of illness at the very beginning. In addition, families in straitened circumstances may not seek medical advice unless emergencies occur because of the cost.

Among the studied systems, the PIM2 scoring system is the most user-friendly, requiring less data collection and only data obtained at the time of admission. The strength of PIM was that the PIM avoids the problem of early treatment bias because it uses only data on admission for prediction. PIM2 has been validated in many countries [6, 8–10, 21, 22], and its performance varies among different counties. In our study PIM2 achieved good calibration and yielded 130.27 predicted deaths, which was similar to the predicted number [130.27 vs. 127, $P = 0.79$ and SMR = 0.97 (95%

CI 0.85–1.12)]. In the subgroup analyses PIM2 performed well in each diagnostic subgroup and medical/surgical subgroup but did not perform well in all of the age subgroups. In the < 1 month subgroup, the PIM2 score underpredicted mortality, with SMR = 1.71 (95% CI 1.18–2.49) and $P=0.04$ ($\chi^2=4.17$, $df=1$). In the adolescent subgroup, the number of predicted deaths was similar to the number of observed deaths (5 vs. 6.52, SMR = 0.77, 95% CI 0.36–1.63), but the discrimination was poor (AUC = 0.66, 95% CI 0.38–0.95). Other validation studies have similarly found that age has an effect on mortality prediction, which is inconsistent with the original PIM study. PIM2 was reported to overpredict death in patients aged > 12 months in a Japanese study and in children aged 1–5 years in an Italian study [8, 10].

The PELOD was a tool to estimate the severity of patients with multiple organ dysfunction syndrome in PICUs and was designed primarily as a surrogate of outcome rather than for prediction of mortality. We evaluated this tool because several recent studies have assessed the performance of the PELOD score as a mortality prediction model and have obtained promising results [4, 5, 23].

Unlike the PRISM and PIM scores, the PELOD score is based on the most abnormal values of variables occurring each day that reflect organ function during the entire PICU stay [3]. This difference may explain why the PELOD-2 score on day one (d1PELOD-2) has achieved excellent predictive performance (AUC = 0.91, 95% CI 0.86–0.96) in septic children [4] because sepsis is defined as the development of organ dysfunction caused by an inappropriate host response to infection [24]. A study among septic children in China also showed PELOD-2 achieved excellent discrimination, although no information on calibration was provided in that study [25]. In our study d1PELOD-2 underestimated mortality (127 vs. 72.9, $\chi^2=205.98$, $P<0.001$) but achieved good discrimination (AUC = 0.80, 95% CI 0.75–0.85) across the whole cohort. The subgroup analyses also showed that the d1PELOD-2 scoring system had acceptable to excellent discrimination but poor calibration. Compared to the present study, in a previous multicenter prospective study, PELOD-2 showed similar discriminatory ability (AUC = 0.80, 95% CI 0.77–0.83) but improved model fitting ($\chi^2=4.81$, $P=0.19$) [20]. The overall mortality and the median (IQR) of PELOD-2 score in that study were both lower than those in our study [4.7% vs. 6.2%, 2 (1–5) vs. 4 (2–5)], which may explain the better calibration in the previous study.

The present study has several limitations. First, the included PICUs were all located in western China, which is less developed than the eastern region of China. The performance of mortality prediction tools may vary among institutions with different medical resources because of corresponding differences in the quality of care in the PICU. Second, there were no NICUs in this study. The bias

may affect the accuracy of the tool in predicting mortality among neonates.

In conclusion, we evaluated the performance of the PRISM, PIM2, PELOD-2, and PRISM IV scoring systems in PICUs in western China. PRISM IV performed best and can be used as a prediction tool for pediatric mortality in PICUs in western China. However, PRISM IV needs to be further validated in NICUs.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12519-022-00603-8>.

Acknowledgements We would like to thank Xiao-Chu Zhong (Imperial College, London, England) for data collection, analysis and interpretation. We thank Yue Zhou (West China Hospital, Chengdu, China), Lu-Ping Wang (West China Hospital, Chengdu, China), Kai-Bo Sun (West China Hospital, Chengdu, China), and Lu-Lu Cao (Mianyang Central Hospital, Mianyang, China) for data collection. We also thank the staff at each PICU for patient enrollment and data collection.

Author contributions XPZ: conceptualization, investigation, formal analysis, writing—original draft, supervision. SYC: conceptualization, writing—review, funding acquisition, supervision. YJ: conceptualization, writing—review, funding acquisition, supervision. YXF: data curation, writing—original draft. YL, GYL, XYZ, XYG, WQZ: data curation. JYZ, HY, GZ, YHH, HW: investigation. CZW, KYY, TQ: formal analysis. All authors approved the final version of the manuscript. Xue-Peng Zhang and Yun-Xia Feng contributed to the work equally.

Funding This work was supported by the National Natural Science Foundation of China (Grant Numbers 81400862 and 81401606), the Key Project of the Science & Technology Program of Sichuan Province (Grant Number 2019YFS0322), the Science Foundation for The Excellent Youth Scholars of Sichuan University (grant number 2015SU04A15), and the 1·3·5 Project for Disciplines of Excellence, West China Hospital of Sichuan University (Grant Numbers 2019HXFH056, 2020HXFH048 and YJC21060).

Data availability The datasets used for the analysis in the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval Ethical approval for the study was provided by the Ethics Committee of the West China Hospital of Sichuan University (NO. 2018–272).

Conflict of interest No financial or non-financial benefits have been received or will be received from any party related directly or indirectly to the subject of this article.

References

1. Pollack MM, Ruttimann UE, Getson PR. Pediatric risk of mortality (PRISM) score. *Crit Care Med*. 1988;16:1110–6.
2. Shann F, Pearson G, Slater A, Wilkinson K. Paediatric index of mortality (PIM): a mortality prediction model for children in intensive care. *Intensiv Care Med*. 1997;23:201–7.

3. Leteurtre S, Martinot A, Duhamel A, Gauvin F, Grandbastien B, Nam TV, et al. Development of a pediatric multiple organ dysfunction score: use of two strategies. *Med Decis Mak.* 1999;19:399–410.
4. Leclerc F, Duhamel A, Deken V, Grandbastien B, Leteurtre S. Can the pediatric logistic organ dysfunction-2 score on day 1 be used in clinical criteria for sepsis in children? *Pediatr Crit Care Med.* 2017;18:758–63.
5. Leteurtre S, Duhamel A, Deken V, Lacroix J, Leclerc F. Daily estimation of the severity of organ dysfunctions in critically ill children by using the PELOD-2 score. *Crit Care.* 2015;19:324.
6. Thukral A, Lodha R, Irshad M, Arora NK. Performance of Pediatric Risk of Mortality (PRISM), Pediatric Index of Mortality (PIM), and PIM2 in a pediatric intensive care unit in a developing country. *Pediatr Crit Care Med.* 2006;7:356–61.
7. Brady AR, Harrison D, Black S, Jones S, Rowan K, Pearson G, et al. Assessment and optimization of mortality prediction tools for admissions to pediatric intensive care in the United Kingdom. *Pediatrics.* 2006;117:e733–42.
8. Imamura T, Nakagawa S, Goldman RD, Fujiwara T. Validation of pediatric index of mortality 2 (PIM2) in a single pediatric intensive care unit in Japan. *Inten Care Med.* 2012;38:649–54.
9. Leteurtre S, Grandbastien B, Leclerc F, Parslow R. International comparison of the performance of the paediatric index of mortality (PIM) 2 score in two national data sets. *Inten Care Med.* 2012;38:1372–80.
10. Wolfler A, Silvani P, Musicco M, Salvo I. Pediatric Index of Mortality 2 score in Italy: a multicenter, prospective, observational study. *Inten Care Med.* 2007;33:1407–13.
11. Eulmesekian PG, Pérez A, Mincos PG, Ferrero H. Validation of pediatric index of mortality 2 (PIM2) in a single pediatric intensive care unit of Argentina. *Pediatr Crit Care Med.* 2007;8:54–7.
12. Slater A, Shann F, Pearson G. PIM2: a revised version of the Paediatric Index of Mortality. *Inten Care Med.* 2003;29:278–85.
13. Pollack MM, Holubkov R, Funai T, Dean JM, Berger JT, Wessel DL, et al. The pediatric risk of mortality score: update 2015. *Pediatr Crit Care Med.* 2016;17:2–9.
14. Murphy-Filkins R, Teres D, Lemeshow S, Hosmer DW. 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;24:1968–73.
15. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression.* New Jersey: Wiley; 2013.
16. Richardson DK, Gray JE, McCormick MC, Workman K, Goldmann DA. Score for neonatal acute physiology: a physiologic severity index for neonatal intensive care. *Pediatrics.* 1993;91:617–23.
17. Taori RN, Lahiri KR, Tullu MS. Performance of PRISM (Pediatric Risk of Mortality) score and PIM (Pediatric Index of Mortality) score in a tertiary care pediatric ICU. *Indian J Pediatr.* 2010;77:267–71.
18. Xu W, Zhang SC. Chinese pediatricians face a crisis: should they stay or leave? *Pediatrics.* 2014;134:1045–7.
19. Coordination Group for National Survey of Development in Pediatric and Neonatal Intensive Care Units. Development of pediatric and neonatal intensive care units: results of a national survey (2000–2009). *Zhonghua er ke za zhi.* 2011;49:669–74 (in Chinese).
20. Zhang Z, Huang X, Wang Y, Li Y, Miao H, Zhang C, et al. Performance of three mortality prediction scores and evaluation of important determinants in eight pediatric intensive care units in China. *Front Pediatr.* 2020;8:522.
21. Czaja AS, Scanlon MC, Kuhn EM, Jeffries HE. Performance of the pediatric index of mortality 2 for pediatric cardiac surgery patients. *Pediatr Crit Care Med.* 2011;12:184–9.
22. Ciofi ML, Cuttini M, Ravà L, Rinaldi S, Brusco C, Cogo P, et al. Performance of the pediatric index of mortality 2 (PIM-2) in cardiac and mixed intensive care units in a tertiary children's referral hospital in Italy. *BMC Pediatr.* 2013;13:100.
23. Karam O, Demaret P, Duhamel A, Shefler A, Spinella PC, Stanworth SJ, et al. Performance of the pediatric logistic organ dysfunction-2 score in critically ill children requiring plasma transfusions. *Ann Inten Care.* 2016;6:98.
24. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA.* 2016;315:801–10.
25. Zhong M, Huang Y, Li T, Xiong L, Lin T, Li M, et al. Day-1 PELOD-2 and day-1 “quick” PELOD-2 scores in children with sepsis in the PICU. *J Pediatr.* 2020;96:660–5.

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.