CrossMark

ORIGINAL ARTICLE

Performance of the pediatric logistic organ dysfunction (PELOD) and (PELOD-2) scores in a pediatric intensive care unit of a developing country

Ahmed El-Nawawy[1] · Aly Abdel Mohsen[1] · Manal Abdel-Malik[1] · Sarah Omar Taman[1]

Abstract The study aimed to compare two scores: the pediatric logistic organ dysfunction (PELOD) with its updated version (PELOD-2) in describing the severity of organ dysfunction in pediatric intensive care unit (PICU) and assess the performance of PELOD-2 in the Egyptian population. A prospective cohort study of 200 patients consecutively admitted to PICU between July 2015 and A 2016 was included. The median age was 6 months, and the male to female ratio was 1.04. The median length of PICU stay was 4 days. The overall predicted number of deaths using PELOD was 76 patients whereas, by PELOD-2, it was 50 patients. The observed mortality was 50 patients. The area under the receiving operating characteristic curve was excellent for both PELOD and PELOD-2 (0.93 and 0.91, respectively). The Hosmer and Lemeshow goodness-of-fit test showed good calibration of PELOD-2 ($\chi^2 = 9.9$, $p = 0.27$), while PELOD showed poor calibration ($\chi^2 = 42$, $p = 0.000$) in the same studied group.

Conclusion: Both scores had excellent discrimination. PELOD-2 is reproducible and easier to perform and had better calibration compared to PELOD score.

What is Known:
• *Pediatric logistic organ dysfunction (PELOD) score was developed 1999 and validated in 2003 to describe the organ dysfunction severity in pediatric intensive care units.*
• *A new and easier version of (PELOD-2) was developed 2013 in France and Belgium to replace the old score. It is important to assess the performance of the new score in other population else than the original.*

What is New:
In an Egyptian pediatric intensive care, the performance of the score revealed:
• *PELOD-2 was an excellent discriminatory score comparable to the original score.*
• *PELOD-2 calibrated well in the Egyptian population while the old score had poor calibration.*

Keywords Pediatric logistic organ dysfunction (PELOD, PELOD-2) · Multiple organ dysfunction syndrome (MODS) · Performance · Validation · Discrimination · Calibration

✉ Manal Abdel-Malik
malakmanal@yahoo.com

Ahmed El-Nawawy
Dr_anawawy@yahoo.com

Aly Abdel Mohsen
aly_elmohsen@hotmail.com

Sarah Omar Taman
sarah.taman@yahoo.com

[1] Department of Pediatrics, Alexandria University, Faculty of Medicine, El-Shatby Hospital Alexandria, Alexandria, Egypt

Abbreviations

AUC	Area under the curve
ICU	Intensive care unit
IQR	Interquartile range
LOS	Length of stay
MODS	Multiple organ dysfunction syndrome
OD	Organ dysfunction
PELOD	Pediatric logistic organ dysfunction
PICU	Pediatric intensive care unit
PIM	Pediatric index of mortality

PRISM   Pediatric risk of mortality
ROC   Receiver operating characteristic
SD   Standard deviation
SMR   Standardized mortality ratio

Introduction

Mortality reduction is an important aim of a pediatric intensive care unit (PICU) [2]. Today's health care environment is focused on providing both high-quality and error-free care. A common adage is "you can't improve what you can't measure." Scoring systems are usually an objective measure which can assess quality of care, assist with the evaluation and modification of complex systems of care, improve patient outcomes, and predict morbidity and mortality [3]. Broadly, these scores can be divided into two categories. The first category belongs to the prognostic scores which predict the risk of death at the time of entry into PICU, and the other category is of the descriptive or outcome scores which describe the course of illness after the admission into PICU [6].

The most frequently used predictive scores in PICUs are Pediatric Risk of Mortality (PRISM) and the Pediatric Index of Mortality (PIM) scores, while the descriptive score widely used to assess multiple organ dysfunction syndrome (MODS) is the Pediatric Logistic Organ Dysfunction score (PELOD) [6].

In 1999, PELOD score was developed using the most abnormal value of each variable during the entire PICU stay [13] and was validated in 2003 [14]. It is by far the most frequently used score aiming to describe the severity of cases of MODS [10]. Because of changes over time in case mix and clinical practice, the performance of this score deteriorated, and there was a need to re-calibrate it. Even though PELOD is quantitative, it is discontinuous, which may cause problems when doing some statistical analyses [18]. In 2013, using a larger and more recent database, the PELOD-2 score was developed and validated with a dataset from two countries: France and Belgium [12].

To be useful, the PELOD-2 should be clinically credible, accurate, and reproducible across other geographic regions. The objective of this study was to evaluate the performance of PELOD and PELOD-2 scores, investigating the relationship between observed outcomes in children admitted in Alexandria University Hospital PICU using both scores in order to substitute PELOD by PELOD-2 in this population.

Subjects and methods

The Alexandria University PICU is nine-bedded, eleven-ventilator unit that admits patients between 1 month and 13 years old. There are three resident doctors on duty each day who answer calls at the "Emergency department," thus minimizing the little delay that exists before ICU admissions. There are three consultants and one assistant on call 24 h; patient nurse ratio is 1:1 round the clock. The number of patients admitted averages 200–250 patients annually.

A sample size of 185 patients was estimated enough required sample to detect a standardized effect size of 0.14 [minimum difference in the area under the Receiver Operating Characteristic ROC curve to detect the primary outcome (mortality) with a minimum required event rate (death) of 18.5% as statistically significant with 90% power and at a significance level of 95% (alpha error = 0.05)]. Sample size increased to 200 patients to control for attrition bias. Sample size was calculated using MedCalc Statistical Software version 14.8.1 [7, 9, 12, 15, 16].

The clinical and laboratory data were collected daily and prospectively recorded on a standardized case report form respecting all aspects of confidentiality. Days were counted by 24-h interval, from the time of admission to 24 h after admission and so on. Routine laboratory tests were taken daily in the morning. Blood gas analysis was measured four times a day and whenever clinically needed. Data collected included age, sex, provisional and final diagnoses, length of stay, variables of both PELOD and PELOD-2 scores, and fate (PICU mortality/discharge).

For the PELOD score, six organ systems (neurologic, cardiovascular, hepatic, respiratory, hematologic, and renal) are considered, each with up to three variables. Each variable is assigned points (0, 1, 10, or 20) based on the level of severity. The maximum number of points for an organ is 20, and the maximum PELOD score is 71 [7].

For the PELOD-2 score, five organ systems (neurologic, cardiovascular, respiratory, renal, and hematologic) are considered and 10 variables were collected. The maximum number of points for an organ is 10, and the maximum PELOD-2 score is 33 [12].

Patients were monitored until death or discharge from the PICU, whichever happened first. Physiologic data from the pre-terminal period (the last 4 h of life) were discarded. If a variable was not measured, we assumed that it was identical to the previous measurement. If a variable was measured more than once in 24 h, the worst value was used in calculating the scores. For each patient, the PELOD and PELOD-2 scores were calculated daily and the worst value recorded during the patient's length of stay in PICU was used for analysis.

Statistical analysis was done using IBM SPSS statistics program version 21 and Medcalc program. A p value ≤0.05 was considered statistically significant. Categorical variables were expressed as frequencies and percentages. Quantitative variables were expressed as median, inter-quartile range (IQR), mean, and standard deviation (SD). Chi-square test or Fisher's exact test was used to study significant association between two categorical variables. Independent sample t and Mann–Whitney tests were used to detect significant difference in the mean, median quantitative variables respectively between two groups of patients. The

choice of each test depends on distribution of variables by Kolmogorov Smirnov test. Z statistic was used to compare the similarity in mortality observed to the expected, through the standardized mortality ratio (SMR) derived from the scores [16].

The discriminant power of the scores was estimated using the area under the receiver operating characteristic curve (AUC) (with 95% confidence interval), the calibration was assessed using the Hosmer–Lemeshow chi-square test, and acceptable calibration is evidenced by a p value ≥ 0.05 [7].

The assessment of consistency or reproducibility of quantitative measurements made by different observers measuring the same quantity was done using the intra-class correlation [15] (using SPSS) and Bland-Altman plot [16] (using MedCalc version 12.2.1.0) was done for agreement.

Results

Among all 225 consecutive patients admitted from July 2015 to April 2016, to the tertiary care PICU of Alexandria University Children's Hospital, 200 patients were included into this prospective observational cohort study. The excluded patients were 25 for the following reasons: incomplete record $n = 13$, PICU stay less than 4 h $n = 3$, children admitted with cardio-respiratory arrest who failed to achieve stable vital signs within 2 h $n = 2$, children admitted for monitoring during elective procedures $n = 2$, patients still cared for $n = 5$.

Table 1 shows characteristics of the patients. PELOD and PELOD-2 scores were significantly higher in non-survivors than in survivors ($p < 0.001$).

Table 2 shows mean PELOD and PELOD-2 and mortality stratified by number of organ dysfunctions. Using the PELOD and PELOD-2 scores, respectively, 13% versus 10% patients had no organ dysfunction (OD) ($n = 26$ versus 20), 22% versus 12% had one OD ($n = 44$ versus 24), 19% versus 32% had two ODs ($n = 38$ versus 64), and 46% versus 46% had >2 ODs ($n = 92$ versus 92). The positive predictive value (PPV) for patient's mortality highest risk (6 ODs) was 100% for PELOD score and 80% for PELOD-2 (5 ODs).

The cutoff value of survival of PELOD score was 13 and the odds ratio for mortality with PELOD score of ≥ 13 was 1.3 (95% CI 1.2–1.4) as compared to the score < 13. On the other hand, the cutoff value of survival of PELOD-2 score was 9 and the odds ratio for mortality with PELOD-2 score of ≥ 9 was 1.5 (95% CI 1.4 to 1.7) as compared to the score < 9. Table 3 shows that PELOD score predicted 76 patients to die among which 48 patients died actually and 28 patients did not, while PELOD-2 predicted 50 patients to die among which 38 patients died actually and 12 patients did not. So, the PELOD sensitivity was significantly higher than PELOD-2 ($P = 0.003$). However, PELOD-2 specificity was significantly higher than PELOD ($P = 0.006$). Both negative and positive predictive values did not show any difference between both scores. AUC values were

0.93 and 0.91 for PELOD and PELOD-2, respectively. The standardized mortality ratio (SMR) using PELOD and PELOD-2 was 0.66 (95% CI 0.49–0.86) and 1.00 (95% CI 0.75–1.30), respectively. PELOD-2 score was found to be significantly closer in prediction of death ($p < 0.001$) compared to PELOD score, and there was no statistically significant difference between both scores as regards their AUC or SMR. The Hosmer and Lemeshow goodness-of-fit test showed a better calibration for PELOD-2 score ($\chi^2 = 9.9$, $p = 0.27$) than for PELOD score ($\chi^2 = 42$, $p = 0.000$) as shown in Table 4.

The Bland-Altman plot showed excellent agreement between PELOD and PELOD-2 scores on the probability of death (Fig. 1). The interclass correlation estimated agreement of 0.897 with a 95% CI (0.844–0.930).

Table 1 Characteristics of studied population admitted to PICU

Baseline characteristics	Value
Gender	
Male, n (%)	102 (51)
Female, n (%)	98 (49)
Age (mo)	
Min-max	1–180
Mean ± SD	20.12 ± 32.01
Median (IQR)	6 (2.5–18.75)
Primary disease at entry, n (%)[a]	
Sepsis	38 (19)
Cardiovascular	18 (9)
Respiratory	88 (44)
Neurologic	18 (9)
Hematologic	20 (10)
Endocrine	12 (6)
Gastrointestinal	44 (22)
Renal	16 (8)
Outcomes	
Mechanical ventilation, n (%)	104 (52)
Length of PICU stay (day), median (IQR)	4 (3–9)
Mortality, n (%)	50 (25)
PIM-2 score	
Mean (SD), median (IQR)	37.57(28.74), 31.1(12–65.4)
PELOD score	
Survived, mean (SD), median (IQR)	9.4 (8.05), 10 (1–12)
Died, mean (SD), median (IQR)	27.5(8.4), 31 (22–32)
PELOD-2 score	
Survived, mean (SD), median (IQR)	4.8 (3.4), 4 (2–6)
Died, mean (SD), median (IQR)	12.8 (4.8), 13 (9.5–17)

SD standard deviation, *IQR* interquartile range, *PICU* pediatric intensive care unit

[a] Patients may have more than one system affected on admission

Table 2 Relation between organ dysfunction, PELOD scores, and mortality

Organ dysfunctions	Patients, n (%)		Mean score (SD)		Mortality, n (%)	
	PELOD	PELOD-2	PELOD	PELOD-2	PELOD	PELOD-2
0	26 (13.0)	20 (10)	0 (0.0)	0 (0)	2 (7.7)	0 (0)
1	44 (22.0)	24 (12)	6.7 (4.4)	2.1 (0.28)	0 (0.0)	2 (8.3)
2	38 (19.0)	64 (32)	9.6 (6.9)	4.7 (1.7)	2 (5.3)	2 (3.1)
3	36 (18.0)	28 (14)	17.6 (7.6)	7.1 (2.2)	8 (22.2)	8 (28.6)
4	24 (12.0)	34 (17)	22.8 (7.1)	11.4 (4.1)	12 (50)	14 (41.2)
5	22 (11.0)	30 (15)	27.2 (4.7)	14 (3.6)	16 (72.7)	24 (80)
6	10 (5.0)		35.0 (7.3)		10 (100)	

PELOD evaluates six organs; PELOD-2 evaluates only five organ dysfunctions

Discussion

In this prospective cohort study, 200 patients admitted to PICU (affiliated to a university teaching hospital which is a tertiary care referral centre providing service to four governorates of nearly 12 million population) were observed and followed up along their stay. The median duration of stay (4 days) was close to data from most PICUs (Portuguese: 3 days, South American: 3 days, French and Belgium: 2 days) [4, 5, 12].

In the present study, 52% of the patients needed the assist of mechanical ventilation (MV). This is very similar to rate of MV in other places of the world. Leteurtre et al. [12] observed MV rate of 52.5%, Garcia et al. [4] reported 35.6%, and Gonçalves et al. [5] found that 68.5% of their patients needing MV.

These scores rely on the association between organ dysfunction, severity of illness, and mortality [11]. PICU mortality is the gold standard against which, we should validate PELOD scores [1]. Among the included 200 patients, mortality rate was 25%. This elevated death rate compared to PICUs of developed countries ranging from (5–16%) [5, 14] could be explained by the higher PIM-2 of the cases admitted to our PICU which means that patients condition was worse on

admission. These patients were admitted on emergency basis in critical condition. While in developed countries PICUs, there is an increased number of electively admitted patients.

In the present study, there was 5.3 and 3.1% of the mortalities having 2 ODs according to PELOD and PELOD-2 respectively which increased to 100% among patients having 6 ODs by PELOD and 80% among patients having 5 ODs by PELOD-2. Leteurtre et al. [14] reported 50% of deceased patients to have 6 ODs. After development of PELOD-2, Leteurtre et al. [12] reported 59% of deceased patients to have 5 ODs. Thus, both PELOD scores are clinically meaningful; the clinical relevance of these scores is significantly linked to mortality in critically ill patients.

In the present study, the PELOD was significantly more sensitive than PELOD-2, while PELOD was statistically less specific compared to PELOD-2 score and both scores did not statistically differ in their predictive values. This means that PELOD score is a better positive test while PELOD-2 is a better negative test for mortality prediction.

The standardized mortality ratio SMR looks at the overall calibration of a score [17]. The equation of probability of death was calculated for all patients using both PELOD and PELOD-2 scores. SMR relates the observed to the expected mortality; if the 95% confidence interval (95% CI) of the SMR includes the value 1.0, the observed number of deaths is not different than the expected number of deaths [11]. In the present study, the 95% CI of the SMR derived from PELOD score was found to be 0.66. In the original validation of PELOD score, SMR was estimated to be 0.55 [14], and Garcia et al. [4] found a SMR of 0.72 using the same score. As regards PELOD-2 score, the present study showed a SMR value of 1.00 compared to 1.42 in its original validation [12] and 1.31 by Gonçalves et al. [5]. The SMR values reported in the present study either from PELOD or PELOD-2 could reflect population difference, in this case the equation needs customization steps [19]. It is clear that PELOD-2 expected deaths value was closer to the observed compare to PELOD in our population but these results were deceiving because PELOD-2 predicted 50 patients to die among which only 76% actually died while 24% did not. PELOD score is a better

Table 3 Discriminate analysis of expected outcome of children admitted to PICU using PELOD and PELOD-2 score

	Observed		Total
	Non-survivors n (%)	Survivors n (%)	
Predicted by PELOD			
Non-survivors	48 (96%)	28 (18.7%)	76 (38%)
Survivors	2 (4%)	122 (81.3%)	124 (62%)
Predicted by PELOD-2			
Non-survivors	38 (76%)	12 (8%)	50 (25%)
Survivors	12 (24%)	138 (92%)	150 (75%)
Total	50 (100%)	150 (100%)	200 (100%)

Table 4 Screening power of PELOD and PELOD-2 scores

	PELOD	PELOD-2	Test of significance
Sensitivity	96%	76%	$Z = 2.8$ ($p = 0.003$)*
Specificity	81%	92%	$Z = 2.7$ ($p = 0.006$)*
PPV	63%	76%	$Z = 1.5$ ($p = 0.12$)
NPV	98%	92%	$Z = 2.4$ ($p = 0.016$)
AUC (95% CI)	0.928 (0.883–0.960)	0.907 (0.858–0.943)	1.781 ($p = 0.075$)[a]
SMR (95% CI)	0.66 (0.49–0.86)	1 (0.75–1.30)	$\chi^2 = 0.0311$ ($p = 0.86$)
Hosmer–Lemeshow test	$\chi^2 = 42$ ($p = 0.000$)*	$\chi^2 = 9.9$ ($p = 0.27$)	NA

PPV positive predictive value, *AUC* area under the ROC curve, *NPV* negative predictive value, *SMR* standardized mortality ratio, *NA* not applicable

*Significance at $p \leq 0.05$

[a] Pair-wise comparison of ROC curves using Delong method (1988)

discriminator of non-survivors 96%. On the other hand, PELOD predicted 124 survivors of which lived 81.3% and PELOD-2 predicted 150 to survive among which 92% actually survived. It is clear now that PELOD expected more deaths compared to PELOD-2 but it was more sensitive in detecting those who will die. This finding was explained by Garcia et al. [4] because PELOD does not recognize important risk-of-mortality intervals (3–16 and 40–80%) and most individuals in these groups are assigned a score which may categorize them in a group above their actual risk-of-mortality.

During the validation process, three aspects of validity need to be addressed: discrimination, calibration, and the clinical utility of the model proposed [4].

Discrimination of these scores was assessed using the area under the ROC curve (AUC), where the result of 1 indicates a perfect discrimination [4]. A discriminatory power of 0.90 or more is considered excellent [5].

In the present study, PELOD and PELOD-2 had an excellent discriminatory power (AUC) of 0.93 and 0.91, respectively, and

the difference was statistically insignificant. These results were similar to many studies. Concerning the PELOD, its original validation reported an AUC of 0.91 [14]. Garcia et al. [4] estimated its AUC to be 0.93. Thukral et al. [17] in a single cohort study in India found the AUC of the PELOD to be 0.8 which is considered to be fair. As regards the PELOD-2 in its original validation set, the value of AUC was 0.98 [12]. Gonçalves et al. [5] assessed the performance of PELOD-2 and found its AUC value to be 0.94 excellent as well.

Calibration The Hosmer–Lemeshow goodness-of-fit test is the most accepted method for measuring calibration. A high p value suggests good calibration and inversely [8].

In the present study, PELOD showed the p value = 0.000 (chi square 42), while PELOD-2 showed a p value = 0.27 (chi square 9.9) using the Hosmer–Lemeshow test. This signifies that PELOD-2 has got a better calibration in the present population than PELOD did in the same studied group of patients. A similar pattern was responsible for the poor calibration of

Fig. 1 Bland-Altman agreement plot on the probability of death
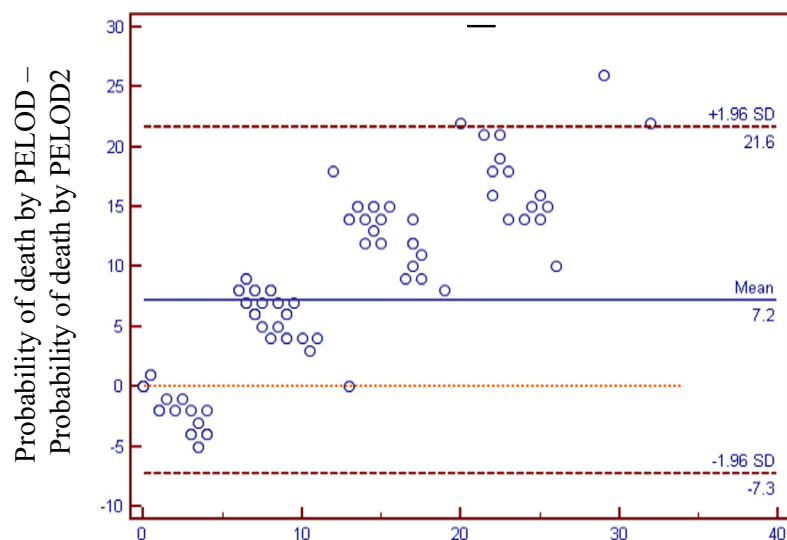
the original validation of PELOD score. The *p* value reported then was <0.01(chi-square 264) [14]. Again, Garcia et al. [4] confirmed this poor calibration with a *p* value less than 0.001 (chi square 72.29). This was compared to a better calibration of PELOD-2; in its original validation, *p* value was 0.317 (chi square 9.31) [12]. These values are very comparable to the present study results.

The reason why PELOD score showed poor calibration compared to PELOD-2 is explained by its under and over prediction of mortality in lower and higher risk groups, respectively. Therefore, PELOD score and its risk-of-mortality are quantitative variables that do not behave as continuous variable which limits its usefulness. While, the updated version PELOD-2 is a continuous scale that can take all integer values [12].

Clinical utility We found that PELOD-2 score was as reliable as PELOD and can be used as surrogate outcome. Both PELOD scores are clinically meaningful, since they can be used as good descriptors of the number and severity of organ dysfunctions in critically ill children during PICU stay independent of the cause. Using the interclass correlation to test for agreement between PELOD and PELOD-2 scores, the estimate was 0.897. This high interclass correlation coefficient was also reported by many other researchers; Garcia et al. [4] estimated 0.87 and Leteurtre et al. [12] found it to be 0.89.

However, PELOD-2 had an advantage over PELOD having fewer variables making assessment more acceptable and also makes the uniform training of PICU staff more convenient [5]. The inter-observer reproducibility was good for all variables of PELOD and PELOD-2 scores because of the use of simple and unequivocal definitions and limited number of variables [14].

This probably is the first study of its kind in African population, so a multicentre study could offer an even better assessment of PELOD-2 performance. Moreover, future researches comparing the worst PELOD-2 during LOS with daily PELOD-2 (d PELOD-2) would provide more strength to the updated score.

Conclusion

This study showed that PELOD-2 score as an update of the PELOD organ dysfunction score allows efficient assessment of the severity of cases of MODS in the PICUs using less variables. PELOD-2 proved to be reproducible has excellent discrimination comparable to that of PELOD score. Moreover, PELOD-2 calibrated well in the present study, whereas PELOD did not. The PELOD-2 score is estimated to be more useful than PELOD score as a surrogate endpoint in clinical trials because it relies on a continuous scale. PELOD-2 score was more specific but less sensitive to predict mortality when compared to PELOD. Meanwhile, it is important to remember that PELOD scores should be used as descriptors of MODS in critically ill patients and not to predict mortality. PELOD-2 score had comparable validation in the present population of a developing country to that of the original validation in developed countries: France and Belgium.

PELOD-2 score is estimated to be an easier and credible score to provide physicians with very useful information, and it is advised to replace the PELOD score with its updated version in PICUs' practice and in clinical researches.

Authors' contributions 1. Ahmed El-Nawawy was responsible for the idea of the research, data analysis, and revision of the manuscript.

2. Aly AbdelMohsen was responsible for revision of the manuscript.

3. Manal Abdel-Malik was responsible for protocol development, data analysis, and writing of the manuscript.

4. *Sarah Omar Taman* was responsible for protocol development and data collection.

References

1. Altman DG, Royston P (2000) What do we mean by validating a prognostic model? Stat Med 19(4):453–473

2. Brady AR, Harrisson D, Black S (2006) Assessment and optimization of mortality prediction tools for admissions to pediatric intensive care in the United Kingdom. Pediatrics 117(4):733–742

3. Buttram SDW, Bakerman PR, Pollack MM (2014). Scoring systems in critical care. Ped Crit Care Med: 47–54.

4. Garcia PCR, Eulmesekian P, Branco RG (2010) External validation of the pediatric logistic organ dysfunction score. Intensive Care Med 36:116–122

5. Gonçalves JP, Severo M, Rocha C (2015) Performance of PRISM III and PELOD-2 scores in a pediatric intensive care unit. Eur J Pediatr 10:2533–2535

6. Gulla KM, Sachdev A (2016) Illness severity and organ dysfunction scoring in pediatric intensive care unit. Indian J Crit Care Med 20(1):27–35

7. Hanley JA, Mcneil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29–36

8. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S (1997) A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 16:965–980

9. Killeen PR (2005) An alternative to null-hypothesis significance tests. Psychol Sci 16(5):345–353

10. Lacroix J, Cotting J (2005) Severity of illness and organ dysfunction scoring in children. Pediatr Crit Care Med 6:126–134

11. Leteurtre S, Duhamel A, Grandbastien B (2010) Daily estimation of the severity of multiple organ dysfunction syndrome in critically ill children. Canad Med Assoc J 182(11):1181–1187

12. Leteurtre S, Duhamel A, Salleron J (2013) PELOD-2: an update of the pediatric logistic organ dysfunction score. Crit Care Med 41: 1761–1773

13. Leteurtre S, Martinot A, Duhamel A (1999) Development of a pediatric multiple organ dysfunction score: use of two strategies. Med Decis Making 19:399–410

14. Leteurtre S, Martinot A, Duhamel A (2003) Validation of the pediatric logistic organ dysfunction (PELOD) score: prospective, observational, multicentre study. Lancet 362:192–197

15. Pannucci CJ, Wilkins EG (2010) Identifying and avoiding bias in research. Plast Reconstr Surg 126(2):619–625

16. Schoonjans F, Zalta A, Depuydt CE (1995) MedCalc: a new computer program for medical statistics. Comput Methods Prog Biomed 48:257–262

17. Thukral A, Khohli U, Lodha R, Kabra SK, Kabra NK (2007) Validation of the PELOD score for multiple organ dysfunction in children. Indian Pediatr 44:683–686

18. Tibby SM (2010) Does PELOD measure organ dysfunction and is organ dysfunction a valid surrogate for death? Intensive Care Med 36:4–7

19. Zimmerman JK, Kramer AA, McNair DS (2006) Acute physiology and chronic health evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med 34: 1297–1310