Validation of Pediatric Index of Mortality-2 Scoring System in a Pediatric Intensive Care Unit, Barbados

by Seetharaman Hariharan,1 Kandamaran Krishnamurthy,2 and Dionne Grannum3

1Department of Clinical Surgical Sciences, The University of the West Indies, St. Augustine, Trinidad & Tobago
2Department of Paediatrics, The University of the West Indies, Cave Hill Campus, Barbados
3Department of Paediatrics, Queen Elizabeth Hospital, Barbados

Correspondence: Seetharaman Hariharan, Department of Clinical Surgical Sciences, The University of the West Indies, St. Augustine, Trinidad & Tobago. Tel/Fax: +1 868 662 4030; E-mail: uwi.hariharan@gmail.com

Summary

This study evaluated the outcome of patients in a pediatric intensive care unit (ICU) of a developing country applying Pediatric Index of Mortality (PIM) version-2 scoring system. A total of 163 consecutive patients were prospectively studied. Data included demographics, diagnoses at admission, PIM-2 score, the duration of ICU stay and hospital outcome. Predicted mortality and standardized mortality ratio (SMR) were calculated. Respiratory and neurological illnesses were the main admission diagnoses. The mean length of stay was 5.4 [95% Confidence Intervals (CI): 4–6.9] days. The mean predicted mortality was 6.2% (95% CI: 4.3–8.1); the observed mortality rate was 5.5%, the SMR being 0.89. Hosmer–Lemeshow analysis calibrated PIM-2 for the case mix $\chi^2 = 5.64$ (df = 7), $p = 0.58$. The area under the ROC curve was 0.82 (95% CI: 0.72–0.92) showing a good discriminant function. Performance of the pediatric ICU in Barbados is comparable to that of developed world by risk-adjusted outcome evaluation.

Key words: intensive care unit, performance, developing country.

Introduction

Many illness severity scoring systems are being used for predicting the outcome of patients admitted to intensive care units (ICUs) in terms of morbidity and mortality [1]. Although it is difficult to predict individual outcome of ICU patients accurately, there have been attempts to codify and validate models, which may prognosticate groups of patients having similar presentation of the illness [2]. Scoring systems are primarily being used to predict the general prognosis of patients, but are additionally being used as performance indicators of ICUs [3]. This is usually accomplished by calculating the risk-adjusted mortality for a particular unit and comparing it with that of another [4]. Pediatric Risk of Mortality-version III (PRISM III), is one of the popular scoring systems used in pediatric ICUs throughout the world [5–7]. Pediatric Logistic Organ Dysfunction (PELOD) is a scoring system specifically used in patients with multi-organ system failure in ICUs [8].

Pediatric Index of Mortality (PIM) was introduced by Shann et al. [9] in 1997 to predict outcome in children admitted to ICUs. Like most other prognostic scoring systems, this system was also updated to PIM-2, which has been better than the earlier version in outcome predictability [10]. PIM-2 is an admission score implying that the scoring system is usually applied during the time of admission, before any ICU intervention is undertaken. This is different to other scoring systems, which could be applied at different time-intervals throughout the ICU stay.

International comparisons of intensive care outcomes are important because critical care delivery patterns and resource consumption may vary considerably in various regions [11]. This is especially true in many developing nations. There has been a many studies comparing the performance of prognostic scoring systems in the pediatric ICUs of different regions [12–14]; but there is still a need for more work comparing different regions of the world [15].

With this background, we conducted this prospective study in a pediatric ICU, to determine the value of PIM-2 scoring system in predicting the outcome of our case-mix, and also assessing the performance of the unit to compare it with international reports.

Methods

Hospital and ICU setting

Barbados is an island nation of the English-speaking Caribbean, with a population of 250,000. The Queen Elizabeth Hospital, Barbados, is a 500-bed tertiary...
care centre, affiliated to the University of the West Indies. The pediatric ICU in QEH is a 4-bed open unit, admitting pediatric patients predominantly belonging to medical specialty but also a few surgical cases. Patients get admitted from Emergency Department directly and from the general wards. A Senior Registrar in pediatric medicine takes care of the unit around the clock under the supervision of a Consultant. The nurse–patient ratio is 1:1. The Radiology department and the Pathology, Microbiology laboratories of the hospital have state-of-the-art equipment facilitating a wide variety of investigations. Majority of the patients admitted to the unit have invasive lines, which include arterial and central venous lines. The unit has facilities for blood gas analyses, portable radiograph and ultrasound.

Approval of Ethics Committee was obtained prior to conducting the study. All patients consecutively admitted to the pediatric ICU over a period of 1 year from January 2008 to December 2008 were enrolled for prospective collection of data. The demographic data recorded were the age and gender of the patients. The diagnoses on admission were noted. PIM-2 scoring system was applied on the day of admission to all the patients before any therapeutic intervention was undertaken. Patients were followed up throughout their stay in ICU and during the entire hospital stay to record their final outcome. The length of stay (LOS) in ICU was recorded. The hospital outcome was classified into either ‘Discharged’ or ‘Died’.

The regression equation published with PIM-2 scoring system was used to calculate the predicted mortality in the ICU patients.

Mann–Whitney U-test was used to compare PIM-2 scores and LOS between survivors and non-survivors. Hosmer–Lemeshow goodness-of-fit analysis was done to calibrate the scoring system. Receiver Operating Characteristic curve (ROC) analyses were done to analyze the discriminant function of the system. Standardized mortality ratio (SMR) (ratio of the observed to the predicted mortality rate) was obtained for the case mix. The statistical significance was fixed at \( p < 0.05 \) level. Statistical analyses were done using the Statistical Package for Social Sciences (SPSS\textsuperscript{®}), (version 12) (Chicago, IL, USA) software.

Results

During the period of study, 163 patients were admitted to the ICU. The median age of the patients was 5 years, 1–8 [Interquartile ranges, (IQR)]. Of these 163 patients, 94 were males (57.7%). Respiratory illnesses such as bronchiolitis and pneumonia were the most common diagnoses on admission followed by neurological and cardiovascular illnesses. The denominations of the diagnostic categories are shown in Fig. 1.

Overall, the patients stayed in the ICU ranging from 1 to 92 days and the mean LOS was 5.4 [95% Confidence Intervals (CI): 4–6.9] days. 84% of the patients had mechanical ventilatory support during their stay in ICU.

The predicted mortality according to the PIM2 score ranged from 0 to 87.8% with a mean of 6.2% (95% CI: 4.3–8.1). The observed mortality rate was 5.5%, the SMR being 0.89. Table 1 shows the age, LOS and the outcomes in all the patients studied. Hosmer–Lemeshow Chi-square analyses for goodness-of-fit for PIM-2 system showed a good calibration of the model to our case-mix [H–L \( \chi^2 = 5.64 \) (df = 7), \( p = 0.58 \)]. The distribution of the deciles of risk in our case-mix has been depicted in Table 2.

Figure 2 shows the ROC curve for PIM-2 scoring system. The area under the curve (AUC) for the ROC curve for PIM-2 system was 0.82 (95% CI: 0.72–0.92).

Discussion

The major finding of the present study is the reasonably good performance of the PIM-2 scoring system in a pediatric ICU of a Caribbean developing country.

There have been many published reports comparing several different versions of pediatric prognostic scoring systems such as PIM, PRISM and PELOD [5, 7, 8]. Some reports found PIM to be performing well in pediatric ICUs compared to other systems.
PIM is also better in terms of its simplicity, validity, discriminatory power, etc. Hence, we applied the current version of PIM in our unit.

The overall age range, diagnostic categories and observed mortality in the present study was mostly comparable to those reported from many parts of the world [16–18]. A study from Trinidad, a neighboring Caribbean country reported that neonates were also admitted in the multidisciplinary ICU because of the policy of the neonatal ICUs, which do not usually admit neonates who had been delivered elsewhere and/or transferred from other neonatal units [18]. However, the pediatric ICU in Barbados does not admit neonates.

The LOS of patients varied widely, although there was no statistical significant difference between survivors and non-survivors. LOS in pediatric ICU has been found to be an important predictor of quality, enabling the comparison of resource utilization among different institutions [19]. Organizational factors known to foster team-oriented care are associated with shorter LOS. Also the size of the pediatric ICU is an important factor in that a relatively larger size pediatric ICU may keep pediatric ICU beds occupied longer [19]. Our pediatric ICU is 4 bedded that would have contributed to the relatively shorter mean LOS. Despite this, there were patients who stayed for months, due to the unavailability of a high-dependency unit in the hospital.

### Table 1
Comparison of survivors and non-survivors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall (n = 163)</th>
<th>Survivors (n = 154)</th>
<th>Non-survivors (n = 9)</th>
<th>Significance*, p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD (years)</td>
<td>4.9 ± 4.4</td>
<td>4.9 ± 4.3</td>
<td>4.7 ± 5.0</td>
<td>0.43</td>
</tr>
<tr>
<td>Gender (n) (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>94 (57.7)</td>
<td>90 (58.4)</td>
<td>4 (44.4)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>69 (42.3)</td>
<td>64 (41.6)</td>
<td>5 (55.6)</td>
<td></td>
</tr>
<tr>
<td>LOS, mean ± SD</td>
<td>5.4 ± 9.2</td>
<td>5.2 ± 8.7</td>
<td>9.1 ± 14.9</td>
<td>0.32</td>
</tr>
<tr>
<td>Predicted mortality, mean ± SD (%)</td>
<td>6.2 ± 12.1</td>
<td>5.4 ± 10.3</td>
<td>20 ± 26.5</td>
<td>0.001</td>
</tr>
</tbody>
</table>

*By Mann–Whitney U-test.

### Table 2
Hosmer-Lemeshow goodness-of-fit analysis for PIM 2

<table>
<thead>
<tr>
<th>Groups (Predicted mortality)</th>
<th>Survivors</th>
<th>Non-survivors</th>
<th>Total (n = 163)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed (n = 154)</td>
<td>Expected</td>
<td>Observed (n = 9)</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
<td>16.4</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>27.9</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>11.6</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>18.3</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>15.4</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>15.3</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>15.2</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>15.1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>18.7</td>
<td>4</td>
</tr>
</tbody>
</table>

Hosmer–Lemeshow $\chi^2 = 5.64$ (df = 7), $p = 0.58$.

The overall age range, diagnostic categories and observed mortality in the present study was mostly comparable to those reported from many parts of the world [16–18]. A study from Trinidad, a neighboring Caribbean country reported that neonates were also admitted in the multidisciplinary ICU because of the policy of the neonatal ICUs, which do not usually admit neonates who had been delivered elsewhere and/or transferred from other neonatal units [18]. However, the pediatric ICU in Barbados does not admit neonates.

The LOS of patients varied widely, although there was no statistical significant difference between survivors and non-survivors. LOS in pediatric ICU has been found to be an important predictor of quality, enabling the comparison of resource utilization among different institutions [19]. Organizational factors known to foster team-oriented care are associated with shorter LOS. Also the size of the pediatric ICU is an important factor in that a relatively larger size pediatric ICU may keep pediatric ICU beds occupied longer [19]. Our pediatric ICU is 4 bedded that would have contributed to the relatively shorter mean LOS. Despite this, there were patients who stayed for months, due to the unavailability of a high-dependency unit in the hospital.

---

**Fig. 2.** ROC curve for PIM-2.
Calibration of a prognostic model is done by testing ‘the degree of correspondence between the probabilities of the outcome as predicted by the score and the observed frequency of the outcome, at different levels of probability’. A good calibration is traditionally represented by a $p$-value $> 0.1$ [20]. Calibration of a prognostic model is usually done by the Hosmer–Lemeshow goodness-of-fit test, which compares the predicted and observed outcomes in subgroups of patients belonging to deciles of risk. The PIM-2 scoring system when introduced, calibrated well for risk deciles but poorly for diagnostic groups. The authors who designed the PIM-2 scoring system have acknowledged the fact that the system may not perform well in different environments. Although there may be a need of changing the coefficients used for the regression equation to suit individual needs, the authors have opined against this alteration, since they feel it may defeat the purpose of the model [6].

There have been many reports from different parts of the world calibrating PIM2. A report from Netherlands calibrated the model and found to be acceptable [$\chi^2 = 4.92$ (df = 8), $p = 0.77$] [12]. A study from Pakistan applied PIM2 and found that it calibrated well [$\chi^2 = 9.65$, $p = 0.29$] [14]. The study from Trinidad had $\chi^2 = 5.61$ (df = 8), $p = 0.69$, pointing to good calibration [18]. Also in Spain, the model calibrated well [$\chi^2 = 4.87$ (df = 8), $p = 0.85$] [21]. The present study also found that PIM2 calibrated well to the case-mix of Barbados.

Discriminant analysis of a model determines its ability to categorize patients into two outcome groups such as survivors and non-survivors. ROC curve analysis is the most commonly applied methodology to determine the discriminatory ability of a prognostic scoring system. Conventionally, an AUC of $>0.7$ is considered acceptable, $>0.8$ is good and $>0.9$ is excellent. The AUC for the ROC for PIM2 has been reported from many countries. It was 0.9 from Australia and New Zealand [8], 0.88 from Pakistan, another developing country, which was similar to a report from Spain (0.87) [14, 21]; our AUC is similar to that reported from India (0.81) [10]. The AUC of the ROC curve reported from Trinidad was quite low (0.62) [18].

The SMR is a well-recognized and valid measure for comparing the risk-adjusted mortality between different centers, although it may vary between regions according to the case-mix as well as the care offered. Additionally, the SMR may be useful as a comparative measure between different scoring systems applied in the same setting. The study from Australia and New Zealand reported an SMR of 0.97 [8] and the SMRs reported from developing countries such as India and Pakistan were 1.57 and 1.4, respectively [10, 14], implying underprediction of mortality. The SMR of the present study is 0.89 was similar to that reported from Trinidad (0.86), which could be considered better than that reported from developed countries [18]. Although this may be interpreted that the unit is performing well, it may be difficult to arrive at such conclusion because like the Trinidad study, there could have been overprediction of mortality. There is an opinion that performance evaluation of ICUs by prognostic models may have errors since many units evaluated this way qualify with ‘honors’ due to ‘grade inflation’ [22].

Our study has few limitations. The number of patients studied may be relatively smaller to calibrate applying goodness-of-fit analysis, although this has been done in many other studies with similar sample sizes. Outcome evaluation in critically ill pediatric patients is a challenging task. It has been well-recognized that this area has not been addressed adequately, which needs plenty of research in the future [23]. Rating the quality of ICUs applying a severity-of-illness scoring system is also highly controversial because performance of a unit is multidimensional [24, 25].

Nevertheless, the present study found that PIM2 scoring system calibrated well and had a reasonable discriminatory ability when applied to the case-mix in a PICU in Barbados. Thus, it could be used as a beneficial tool for evaluation of risk-adjusted mortality and comparison of units within the region.

**References**


22. Popovich MJ. If most intensive care units are graduating with honors, is it genuine quality or grade inflation? Crit Care Med 2002;30:2145–6.

