The PATRIARCH* Study. Using outcome measures for league tables: Can a North American prediction of admission score be used in a United Kingdom children’s emergency department?

H Miles, E Litton, A Curran, L Goldsworthy, P Sharples, A J Henderson

Objectives: The use of league tables has become predominant in the healthcare culture of the United Kingdom. These tables are often based on measures that are viewed with scepticism by clinicians. This study was designed to test the validity of a North American risk of admission score, the PRISA, for use in a United Kingdom population of accident and emergency (A&E) attendees.

Methods: All attendees to a children’s A&E department were scored using the PRISA for a single calendar month (November 2000).

Results: 701 children were studied in total. The results show that the PRISA applied to this population gives an area under the receiver operator curve of 0.76. Of the 701 patients studied, 206 (29.4%) were admitted. The PRISA predicted a total of 206.10 admissions. Of the 50 patients discharged with the highest PRISA scores (that is, with the highest likelihood of admission), none were admitted in the 48 hours after their original attendance.

Conclusions: These results show that the PRISA is suitable as a measure of paediatric A&E department performance in the United Kingdom and it is highly promising as a future measure of quality.

METHODS

Setting
The Accident and Emergency Department at Bristol Royal Hospital for Sick Children is part of a tertiary referral hospital and receives about 15,000 patients a year, 12,500 (83%) of whom are first time attendees. The department is staffed by one full time consultant, two registrars, and five senior house officers.

Population
Patients come from the greater Bristol area and attend the department by self referral, general practitioner referral, or direct ambulance transport. The population of Bristol is predominantly white, 94.9% white, 2.1% coloured, 2.4% black, and 0.6% other. During a single calendar month (November 2000), all attendees to the accident and emergency (A&E) department were scored on the PRISA.

Data collection and analysis
The outcome of the visit was recorded (admission or non-admission to a hospital bed) as well as details of category of illness and final clinical diagnosis. Clinical measurements were made and recorded by medical and nursing staff in the department who had been briefed on the study and the PRISA score at the outset. This briefing entailed the placing of posters around the department giving a start date, the clinicians...
involved, and referring readers to the PATRIARCH questionnaire that was placed in the unit communication book.

The data collected were coded by the investigators (HM, EL) and transformed to risk of admission probabilities using the formulas described by Chamberlain and others. Children with minor injuries only were excluded from analysis as the PRISA score is not validated for use in this population.

The outcome of each emergency attendance was recorded as admission or non-admission. We also examined the non-admissions and selected the 50 children with the highest probability of admission from this population. Their details were cross checked against the hospital admissions records to determine whether they were subsequently admitted during the 48 hours after their initial attendance.

The goodness of fit of the PRISA score to the observed outcomes in our population was examined by two methods: (1) the performance of PRISA in discriminating between admissions and non-admissions was tested using the area under the receiver operator characteristic curve; (2) the calibration was tested by comparing observed with predicted admissions and non-admissions across quintiles of risk and using the Hosmer and Lemeshow goodness of fit test.

Descriptive analyses of the population attending our A&E department used non-parametric tests ($\chi^2$ and Kruskal-Wallis analysis of variance). Data are presented with 95% confidence intervals and a level of statistical significance of $p<0.05$ was used throughout. Analyses were performed using SPSS v9.0 for Windows.

**RESULTS**

Information was collected on 944 consecutive A&E attendances of whom 243 were excluded as having minor injuries only. On checking departmental attendance records, we discovered a further six patients who had attended the department but for whom no data were collected. Our final analysis was performed on 701 children attending the paediatric A&E department. Their ages ranged from 0–20 years (median (interquartile range) 2 years (1–7 years)) and the male/female ratio was 1.30 (396 boys; 305 girls). The source of referral was the general practitioner in 264 cases (37.7%), direct ambulance in 95 cases (13.6%), walk in patients (self referral) in 301 cases (42.9%), and other (for example, referral from another hospital or unit) in 41 cases (5.8%).

**Diagnoses (Table 1)**

A total of 15 diagnostic groups were identified. The three commonest were respiratory (18.8%), surgical (17.8%), and gastrointestinal (17.1%). The least common medical condition was allergy (0.4%).

**Goodness of fit of PRISA**

The receiver operator curve, plotting true positive rate (sensitivity) against false positive rate (1−specificity) at different cut off points of PRISA, is shown in figure 1. The area under the curve (95% CI) was 0.76 (0.72 to 0.80). A perfect test would have an area of 1.0 and one relying completely on chance would have an area of 0.5. Table 2 shows the observed and predicted admissions across quintiles of risk based on PRISA probabilities of admission ($\leq20\%$; $>20\%\leq40\%$; $>40\%\leq60\%$; $>60\%\leq80\%$, and $>80\%$ risk of admission). The Hosmer and Lemeshow goodness

**Table 2**

<table>
<thead>
<tr>
<th>Risk of admission</th>
<th>$\leq20%$ (n=251)</th>
<th>$&gt;20%\leq40%$ (n=300)</th>
<th>$&gt;40%\leq60%$ (n=78)</th>
<th>$&gt;60%\leq80%$ (n=58)</th>
<th>$&gt;80%$ (n=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admitted</td>
<td>33</td>
<td>78</td>
<td>44</td>
<td>39</td>
<td>12</td>
</tr>
<tr>
<td>Predicted</td>
<td>28.00</td>
<td>87.57</td>
<td>39.04</td>
<td>39.29</td>
<td>12.20</td>
</tr>
<tr>
<td>O:E*</td>
<td>1.18</td>
<td>0.89</td>
<td>1.13</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>95%CI</td>
<td>1.15 to 1.20</td>
<td>0.88 to 0.91</td>
<td>1.10 to 1.15</td>
<td>0.97 to 1.02</td>
<td>0.95 to 1.01</td>
</tr>
<tr>
<td>Discharged</td>
<td>218</td>
<td>222</td>
<td>34</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>Predicted</td>
<td>223.00</td>
<td>212.43</td>
<td>38.96</td>
<td>18.71</td>
<td>1.80</td>
</tr>
<tr>
<td>O:E*</td>
<td>0.98</td>
<td>1.05</td>
<td>0.87</td>
<td>1.02</td>
<td>1.11</td>
</tr>
<tr>
<td>95%CI</td>
<td>0.98 to 0.98</td>
<td>1.04 to 1.05</td>
<td>0.85 to 0.89</td>
<td>0.97 to 1.07</td>
<td>0.92 to 1.32</td>
</tr>
<tr>
<td>O:E; observed:expected ratio (standardised admission ratio).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3**

<table>
<thead>
<tr>
<th>GP referrals</th>
<th>Ambulance</th>
<th>Walk in</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRISA score median [IQR]</td>
<td>10 [10–17.75]</td>
<td>11 [8–18]</td>
<td>0 [0–8]</td>
</tr>
<tr>
<td>Male/female ratio</td>
<td>1.64</td>
<td>1.11</td>
<td>1.18</td>
</tr>
<tr>
<td>Discharges</td>
<td>167</td>
<td>51</td>
<td>250</td>
</tr>
<tr>
<td>Admissions</td>
<td>97</td>
<td>44</td>
<td>51</td>
</tr>
<tr>
<td>Predicted admissions</td>
<td>105.92</td>
<td>39.83</td>
<td>49.15</td>
</tr>
<tr>
<td>95% CI</td>
<td>100.43 to 111.41</td>
<td>36.05 to 43.61</td>
<td>43.67 to 52.63</td>
</tr>
<tr>
<td>95%CI</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
</tr>
</tbody>
</table>

O:E; observed expected ratio (standardised admission ratio).
of fit test did not demonstrate a statistically significant deviation of observed from predicted admissions and discharges across these risk intervals ($H=3.34, p=0.3$).

**Analysis of local population**

Of the 701 patients that we studied, 206 (29.4%) were admitted to a hospital bed and 495 were discharged from the A&E department. Of the 50 patients discharged with the highest PRISA scores, none was admitted in the 48 hours after their initial attendance. PRISA predicted a total of 206.10 admissions, resulting in a standardised admission ratio (95%CI) of 1.00 (0.95 to 1.05).

Table 3 shows the number of admissions and discharges by mode of referral. There were no statistically significant differences in the ages of children (Kruskal-Wallis, $p>0.1$) or proportions of boys ($\chi^2=0.01$) referred from each source but walk in patients had statistically significantly lower PRISA scores than patients referred from other sources (Kruskal-Wallis, $p<0.001$). The proportion of children admitted from self referred (walk in) patients was also statistically significantly lower than from other sources ($\chi^2=43.42, p<0.001$).

**DISCUSSION**

Our results show that the PRISA is suitable as a measure of paediatric A&E department performance in the United Kingdom. Although it was only able to predict outcome with 70% accuracy this level of accuracy is the same as that found in North America. It is possible that collecting the data in November may have introduced a bias in the results, as there may have been a preponderance of respiratory illness. We did not find this to be the case. It is also interesting to note that the populations studied had a widely different ethnic mix, the North American population being predominantly black, while the population studied in the present work was predominantly white. It would seem therefore that the PRISA can be used in a United Kingdom population and it is highly promising as a future measure of quality.

The PRISA is a carefully developed and reasonably sophisticated score relying on a detailed evaluation of the patient (appendix 1, see journal web site; emjonline.com). On the surface therefore, it should have provided a better degree of accuracy. Its inability to do so raises again the issue regarding the use of much cruder outcome measures such as mortality rates to assess hospital performance. A parallel has been drawn between the use of league tables in the health service and those used in education by Fletcher. He concludes that most of the present outcome measures in both health and education are poor indicators of quality, and that social and economic factors are much more relevant to the outcomes measured than the work of individual institutions.

The use of league tables by the United Kingdom Department of Health to compare hospital performance also rely on outcome measures that have been heavily criticised on methodological grounds, on their inability to provide valid comparisons, on questions regarding patient case mix selection, on their reliability as indicators of performance or quality of care, and on the inability of the various parties to agree as to how performance should be measured. Despite being somewhat revised in 1999, these criticisms have not yet been fully met. The league tables themselves, by focusing on impersonal and averaged figures shift care away from a person centred approach and constrain the possibilities of responsiveness.

If league tables are to continue to be published, as seems probable, it will be necessary to find measures of outcome that answer these criticisms. One such criticism is the poor comparability between outcomes when illness severity is not considered in the calculations. This is not to say however that measuring performance is not important. To do this effectively measures must be found that satisfy parties on both sides of the clinician/management divide that are meaningful. It is only by achieving this that the individual clinicians on whom the health service depends to treat patients will have confidence in league tables and be able to see them as something more than the desire of government to one sidedly challenge the claims of professionalism adopted by those who work in the health service.

**ACKNOWLEDGEMENTS**

The authors wish to acknowledge the support of the nursing and medical staff of the Accident and Emergency Department, Bristol Royal Hospital for Sick Children in conducting this study.

**Contributors**

Helen Miles initiated the original idea, discussed core ideas, designed the study questionnaire, participated in protocol design, participated in data collection, analysis and reading of the paper. Edward Litton initiated the original idea, discussed core ideas, designed the study questionnaire, participated in data collection, analysis and reading of the paper. Andrew Curran coordinated the formulation of the primary study hypothesis, discussed core ideas, participated in protocol design, analysis and interpretation of the data and the writing of the paper. Lisa Goldsworthy discussed core ideas, participated in protocol design, analysis and interpretation of the data and the writing of the paper. Peta Sharples discussed core ideas and participated in protocol design. John Henderson discussed core ideas, participated in protocol design, performed the statistical analysis and participated in the interpretation of the data and the writing of the paper. Andrew Curran and John Henderson will act as guarantors of this paper.

**Authors’ affiliations**

H Miles, E Litton, A J Henderson, University of Bristol, UK
A Curran, Frenchay Hospital, Bristol, UK
L Goldsworthy, P Sharples, Bristol Children’s Hospital, UK

**REFERENCES**

10 Fletcher J. Encouraging health care policy: are league tables likely to dig out the truth about hospital performance or dig it in? Journal of Nursing Management 1995;5:255–62.