Surgical audit, in some form or the other, has been practised since long, but it has been accepted in regular practice only recently. Surgical audit consists of critical assessment of surgical outcome and results. There are many varieties of audits but all require critical assessment of surgical outcome. If the risk of an adverse outcome is known for a group of patients, the actual outcome can be compared with the predicted outcome, and comparison can be made between groups in different surgical units for the purposes of audit or research. The pressing need to develop measures of health outcome for use in surgical audit was recognized and resulted in the development of POSSUM, a Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity, a scoring system developed to provide risk-adjusted analysis in patients undergoing surgery; but have not been validated in low-risk patients. The aim of this study was to assess POSSUM and P-POSSUM in general surgical patients in a developing country. Materials and Methods: Seven hundred and eighty-eight consecutive general surgical patients were studied prospectively with POSSUM and P-POSSUM scoring systems using linear and exponential analysis. The ratio of observed to predicted death and morbidity (O: E) was calculated for each analysis and frequency tables were compared for statistical significance by means of chi square test. In the second part of this study prospective evaluation of Modified POSSUM was performed on the next 908 patients. Results: POSSUM was found to be a good predictor of mortality and morbidity with exponential analysis but not with linear analysis. P-POSSUM predicted the mortality accurately with the help of linear analysis. Both overpredicted the outcome in the low-risk group (predicted risk of mortality < 10% and predicted risk of morbidity < 40%); therefore POSSUM needs to be modified for audit in the low-risk group. Such a modification, Jabalpur-POSSUM (J-POSSUM: a modification of POSSUM for low-risk group) revealed good fitness with observed mortality (O: E = 1.00) and morbidity (O: E = 1.04). On prospective evaluation of J-POSSUM, this finding was confirmed for the low-risk group vis-à-vis observed mortality (O: E = 1.16) and morbidity (O: E = 0.83). Conclusion: POSSUM is not a good predictor of low-risk patients and needs risk adjustment with the help of correcting factor for accurately predicting the mortality and morbidity. Such a correction factor was identified with the help of multiple logistic regression analysis and prospectively validated. Jabalpur-POSSUM can successfully predict the outcome in low-risk patients.

Key Words: Surgical audit, POSSUM, morbidity, mortality, risk-adjusted analysis

Original Article


Free full text available from http://www.indianjsurg.com
system which assesses perioperative surgical risks. This scoring system produced assessments for morbidity and mortality rates, which did not significantly differ from observed rates and has been acknowledged as the most appropriate of the currently available scores for general surgical practice. Although POSSUM and its Portsmouth modification (P-POSSUM) have been successfully validated in many different patient subsets, they have never been validated in low-risk patients in a developing country. The aim of the present study was to prospectively examine the value of POSSUM and P-POSSUM in predicting mortality and morbidity in general surgical patients in a developing country, failing which to develop a modification which works best for this group of patients.

MATERIALS AND METHODS

The first part of this prospective study was carried out in the Department of Surgery, NSCB Government Medical College, Jabalpur (MP), INDIA during October 2000 to September 2001 on 788 consecutive adult general surgical patients undergoing elective or emergency surgery and requiring in-patient care of at least 24 hours after operation. POSSUM score [consisting of two categories of assessment to assess the risk of surgery: a 12-factor (age, cardiac status, pulse rate, systolic blood pressure, respiratory status, Glasgow Coma Score, serum concentration of urea, potassium and sodium, haemoglobin concentration, white cell count and findings on electrocardiography) and 4-grade physiological score (PS) combined with a 6-factor (type of surgical procedure, number of procedures, blood loss, peritoneal soiling, presence of malignancy and mode of surgery) and 4-grade operative severity score (OSS).] was calculated and morbidity and mortality were recorded as defined by Copeland et al. Complications were assessed by clinical observation. Routine bacteriological screening and postoperative radiological scanning were not carried out but confirmatory bacteriological and radiological tests were carried out where clinical doubts existed.

Using outcome (dead/alive or complicated/uncomplicated) as a dichotomous dependent variable, multiple logistic regression equation derived by POSSUM was applied to all patients vis-à-vis both morbidity and mortality. Equation for morbidity was:

\[
\text{In } \log \left( \frac{R}{1-R} \right) = -5.91 + (0.16 \times \text{physiological score}) + (0.19 \times \text{operative severity score})
\]

For mortality the equation used was:

\[
\text{In } \log \left( \frac{R}{1-R} \right) = -7.04 + (0.13 \times \text{physiological score}) + 0.16 \times \text{operative severity score}
\]

[Where R is the predicted risk]

Additionally, P-POSSUM equation was applied for mortality as follows:

\[
\text{In } \log \left( \frac{R}{1-R} \right) = -9.37 + (0.19 \times \text{physiological score}) + 0.15 \times \text{operative severity score}
\]

POSSUM and P-POSSUM equations were applied to all patients and tested for goodness of fit. Linear analysis was done by calculating patient's predicted risk of death using the respective equation and then dividing the patients into groups according to their predicted risk of death. For each version of POSSUM, the number of patients falling into each mortality group was multiplied by the average risk of death to give the predicted number of deaths in that group. Exponential analysis was done by considering a cut-off risk of death in each stage of the calculation, then grouping together all patients whose predicted risk fell above the cut-off point. The ratio of observed to predicted death (O:E) was calculated for each analysis and frequency tables were compared for statistical significance by means of chi square test. The same method was applied for complications.

The second part of this study consisted of prospective evaluation of Modified POSSUM (derived by the first part of this study). This was done during October 2001 to September 2002 on 908 consecutive adult general surgical patients undergoing elective or emergency surgery and requiring in-patient care of at least 24 hours after operation. The ratio of observed to predicted death (O:E) was calculated for low-risk patients and compared for statistical significance by means of chi square test. The same method was applied for complications.

RESULTS

Seven hundred and eighty-eight patients underwent elective or emergency surgery during the first half of the study period, which required in-patient care for at least 24 hours, 594 were male and 194 female. Three hundred and ninety-one patients were operated as elective surgery while 397 patients were operated as emergency surgery. The types of surgical procedure performed were as follows: gastrointestinal 355, hepatobiliary 25, urological 68, hernia 77, breast 69 and miscellaneous 194. The overall mortality rate was 6.72% and the overall morbidity rate was 19.41%. The range of physiological scores obtained is shown in Figure 1 and that of operative severity scores in Figure 2.
There were a total of 53 deaths. Using linear analysis, POSSUM overpredicted the mortality as compared to observed mortality, while using exponential analysis, the predicted and observed mortality rates were similar (Table 1). The mortality rate predicted by P-POSSUM, using linear analysis was as observed but when exponential analysis was used it was significantly higher than observed mortality (Table 1). Postoperative complications seen were haemorrhage (wound: n = 21, deep: n = 5, others: n = 1), infection (respiratory tract: n = 85, wound: n = 140, urinary tract: n = 38, deep: n = 30, sepsicaemia: n = 46, and others: n = 31), wound dehiscence (superficial: n = 85, and deep: n = 68), anastomotic leak (n = 33), thrombosis (deep vein thrombosis: n = 6, pulmonary embolism: n = 3), cerebrovascular accident: n = 1, myocardial infarction: n = 5, and others: n = 4), cardiac failure (n = 13), hypotension (n = 42), respiratory failure (n = 18), renal failure (n = 25) and others (n = 22). The morbidity rate predicted by POSSUM by linear analysis was significantly higher than observed morbidity while in exponential analysis it was as predicted (Table 1).

When POSSUM analysis was done on patients undergoing emergency surgery (n = 397), the O: E ratio was found to be 0.78 for mortality ($X^2 = 2.81$, 1 d.f., $P>0.05$) and 0.86 for morbidity ($X^2 = 3.02$, 1 d.f., $P>0.05$) revealing a good fitness.

In the present study nearly 70% patients (529 out of 788) were in the low-risk group (predicted risk of mortality <10% and predicted risk of morbidity <40%). This group showed very poor fit, with O:E ratio of 0.27 for mortality and 0.65 for morbidity by POSSUM analysis and 0.54 for mortality by P-POSSUM analysis (Table 1). Given this poor fit, the need for a correction factor was identified and multiple logistic regression analysis was done. A correction factor of 0.257 for mortality and 0.619 for morbidity for POSSUM risk group was obtained. After applying this correction factor, Jabalpur POSSUM score (J-POSSUM) for low-risk general surgical patients was obtained.

J-POSSUM was prospectively validated in the next 908 consecutive patients who underwent elective or emergency surgery during the second half of the study period, which required in-patient care for at least 24 hours. Of these, 689 were male and 219 female. Four hundred and sixty patients were operated as elective surgery while 502 patients were operated as emergency surgery. The types of surgical procedures performed were as follows: gastrointestinal 408, hepatobiliary 22, urological 127, hernia 76, breast 40 and miscellaneous 235. The overall mortality rate was 7.38% and the overall morbidity rate was 18.50%. The range of physiological scores and operative severity score was similar to that shown in Figures 1 and 2. J-POSSUM score correlated well with O: E ratio of 1.00 for mortality and 1.04 for morbidity showing no evidence of lack of fit (Table 1). The results of the first half of this study are summarized in Table 1.

### Table 1: Risk adjusted analysis (Summary of first half of study)

<table>
<thead>
<tr>
<th>Type of analysis</th>
<th>Risk group</th>
<th>No. of patients</th>
<th>Mean risk</th>
<th>O</th>
<th>E*</th>
<th>O:E ratio</th>
<th>Level of significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall mortality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSSUM Linear</td>
<td>0-100</td>
<td>788</td>
<td>10.37</td>
<td>53</td>
<td>82</td>
<td>0.86</td>
<td>$X^2=32.38$, 9 d.f. $P&lt;0.001$</td>
</tr>
<tr>
<td>POSSUM Exponential</td>
<td>0-100</td>
<td>788</td>
<td>10.37</td>
<td>53</td>
<td>82</td>
<td>0.94</td>
<td>$X^2=16.22$, 9 d.f. $P&gt;0.05$</td>
</tr>
<tr>
<td>P-POSSUM Linear</td>
<td>0-100</td>
<td>788</td>
<td>4.69</td>
<td>53</td>
<td>37</td>
<td>1.525</td>
<td>$X^2=15.55$, 9 d.f. $P&lt;0.05$</td>
</tr>
<tr>
<td>P-POSSUM Exponential</td>
<td>0-100</td>
<td>788</td>
<td>4.69</td>
<td>53</td>
<td>37</td>
<td>1.42</td>
<td>$X^2=27.46$, 9 d.f. $P&lt;0.05$</td>
</tr>
<tr>
<td><strong>Morbidity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSSUM Linear</td>
<td>0-100</td>
<td>788</td>
<td>34.64</td>
<td>153</td>
<td>273</td>
<td>0.55</td>
<td>$X^2=68.69$, 9 d.f. $P&lt;0.000$</td>
</tr>
<tr>
<td>Exponential</td>
<td>0-100</td>
<td>788</td>
<td>7.92</td>
<td>153</td>
<td>175</td>
<td>0.87</td>
<td>$X^2=17.84$, 9 d.f. $P&lt;0.05$</td>
</tr>
<tr>
<td><strong>Low risk group (mortality)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSSUM Linear</td>
<td>0-10</td>
<td>529</td>
<td>3.36</td>
<td>5</td>
<td>18</td>
<td>0.27</td>
<td>$X^2=9.38$, 1 d.f. $P&lt;0.05$</td>
</tr>
<tr>
<td>P-POSSUM Linear</td>
<td>0-10</td>
<td>714</td>
<td>2.76</td>
<td>12</td>
<td>22</td>
<td>0.54</td>
<td>$X^2=4.55$, 1 d.f. $P&lt;0.05$</td>
</tr>
<tr>
<td>J-POSSUM Linear</td>
<td>0-10</td>
<td>529</td>
<td>0.87</td>
<td>5</td>
<td>5</td>
<td>1.00</td>
<td>$P&gt;0.000$</td>
</tr>
<tr>
<td><strong>Low risk group (morbidity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSSUM Linear</td>
<td>0-40</td>
<td>437</td>
<td>8.00</td>
<td>23</td>
<td>35</td>
<td>0.65</td>
<td>$X^2=4.14$, 1 d.f. $P&lt;0.05$</td>
</tr>
<tr>
<td>J-POSSUM Linear</td>
<td>0-40</td>
<td>437</td>
<td>4.95</td>
<td>23</td>
<td>22</td>
<td>1.04</td>
<td>$X^2=0.04$, 1 d.f. $P&gt;0.05$</td>
</tr>
</tbody>
</table>

O=Observed, E=Expected, *Rounded to nearest whole number.
DISCUSSION

Audit is definitely much more than only data collection, it is complementary to research, education, a commitment to improvement in care by stimulating further analyses, ensuring that practice is recorded, reviewed and made accountable, thereby resulting in improved practice habits.

POSSUM, a popular system of surgical audit has been widely used for comparative audit, comparisons between surgeons, and units, disease groups, and between two scores. P-POSSUM, a modification of POSSUM was developed following reports that POSSUM tends to overestimate the mortality, the P-POSSUM equation producing a very close fit with the observed in-hospital mortality. P-POSSUM, although applied successfully in vascular and gastrointestinal surgery patients, does not estimate morbidity.

In a developing country like India these risk-adjusted evaluations have not been done, perhaps, because of difficulty in the collection of accurate data, differences in patient presentation, follow-up difficulties, and limited financial resources. Testing for goodness of fit with the data, to which it is being applied, is a must for any prognostic scoring system. Geographical variation in the different patient subsets makes such testing and validation mandatory. Since each surgical unit serves a different patient population, each score system must be calibrated in the individual hospital to ensure that the model is applicable for the patient material involved, before the scoring system is accepted as quality standard. This prompted us to attempt the prospective validation of POSSUM and P-POSSUM in our patients.

In our study POSSUM performed well vis-à-vis mortality and morbidity only when exponential analysis was used; the use of linear analysis resulted in overprediction of mortality and morbidity. P-POSSUM predicted the mortality well when linear analysis was used but failed to do so when exponential analysis was used. Wijesinghe et al have tried to explain the propensity of POSSUM to overpredict mortality when inappropriate analysis is used, the O:E ratios for POSSUM being close to unity when the appropriate analysis is performed i.e. exponential analysis.

Exponential analysis is not convenient because it has to be stopped and restarted at a new level if the predicted number of death falls below that calculated at a higher cut-off. The point at which the recalculation has to be performed may vary between populations depending on their spread of predicted risk of death. It is unclear how subgroup analysis can be performed by exponential analysis. A further difficulty with exponential analysis is its need to count and recount patients in the risk band. A patient in the 90-100% band is counted not only in this band but also in the 80-100%, 70-100%, 60-100%, 50-100%, 40-100%, 30-100%, 20-100% and 10-100% bands; whereas a patient whose risk of death is 5% is counted only once, i.e. in 0-100% band. The prediction of death in individual patients may therefore be inaccurate. Notwithstanding these drawbacks the present data show that POSSUM when used with exponential analysis provides a reasonably accurate prediction of death and complications in the whole population studied.

Analysis of morbidity and mortality in the low-risk group revealed that both could not be predicted accurately. POSSUM and P-POSSUM, both overpredicted the mortality in low-risk group, which forms the majority of our patients (Table 1). This is the most important group for audit purposes since it contains the majority of surgical patients and is composed of fit patients undergoing minor surgery. In non-specialist, general surgical units in developing countries with a broad case mix, a reasonable standard of care and practice produces very low mortality and an acceptable morbidity rate. Since death is a relatively rare health outcome in these patients, the development of a more acceptable measure, e.g. morbidity, must be a priority. Any audit system will not be complete without including these patients and their morbidity—an important point against the use of P-POSSUM. The use of multiple logistic regression analysis made it possible to identify the correlation co-efficient, which when used as a correction factor with traditional POSSUM equation, gave an accurate estimate of mortality and morbidity; giving rise to J-POSSUM (Tables 1 and 2).

A key feature of any audit system has to be the capacity both for local audit within hospitals and for global audit between hospitals. POSSUM fulfils the criteria for the most appropriate audit system as it accurately assesses both, morbidity and mortality rates but requires modification by a correction factor for risk adjustment in the low-risk general surgical patients. Such a correction factor was identified with the help of multiple logistic regression analysis, and modified POSSUM (J-POSSUM) can, then, successfully predict the outcome in low-risk patients. Since it can be argued that J-POSSUM was developed by applying a correction

### Table 2: Prospective Analysis of low-risk general surgical patients with J-POSSUM (Summary of the second half of the study)

<table>
<thead>
<tr>
<th>Type of Score</th>
<th>Risk group</th>
<th>No. of Patients</th>
<th>Observed (n)</th>
<th>Expected (n)</th>
<th>O:E ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>0-&lt;10</td>
<td>636</td>
<td>7</td>
<td>6</td>
<td>1.16</td>
</tr>
<tr>
<td>Morbidity</td>
<td>0-&lt;40</td>
<td>585</td>
<td>52</td>
<td>63</td>
<td>0.83</td>
</tr>
</tbody>
</table>
factor based on the results of one particular set of patients and is therefore bound to reveal perfect fitness when applied to this set of data, a prospective evaluation of J-POSSUM was done in a new set of patients. This prospective evaluation has also validated the success of J-POSSUM (Table 2), thereby showing that Jabalpur-POSSUM can successfully predict the outcome in low-risk patients.

REFERENCES