We describe the development and validation of a scoring system for auditing orthopaedic surgery. It is a minor modification of the POSSUM scoring system widely used in general surgery. The orthopaedic POSSUM system which we have developed gives predictions for mortality and morbidity which correlate well with the observed rates in a sample of 2326 orthopaedic operations over a period of 12 months.

Received 11 July 2001; Accepted after revision 17 December 2001

The assessment of outcome after surgical intervention is not a new science. As early as 1750 BC King Hammurabi of Babylon issued a number of decrees relating to surgeons and their surgery. The most infamous of these codices was that if a surgeon operated on a free man and the patient became blind or worse still died, the surgeon should have his operating hand cut off. While, to some a modification of this codex may still seem to be in operation, many have attempted to devise more reliable and robust methods for assessing the outcome of surgical intervention.

1-3 The use of raw data on mortality and morbidity from different units produces disparities in outcome which may be explained by variations in the case mix and in techniques of surgery.4,5 Methods have been devised to allow comparison between units by taking the physiological status of the patients and the operative complexity into account.1,6,7

In general surgery the POSSUM and P-POSSUM systems have proved to be the most reliable and widely applicable scoring methods so far devised.1,8 Both are found to be equally applicable to subspecialties including vascular surgery, surgical gastroenterology and urology and are used by many health-care organisations.7-9

A number of orthopaedic studies have drawn attention to the variability in outcome after surgery, in particular for fracture of the neck of the femur, and it has been argued that this is more likely to be related to the case mix than to the hospital facilities or the surgeons.12,13 A recent report has cautioned that predicting the outcome of orthopaedic intervention in the elderly injured patient, merely on the basis of the injury severity score, can be hazardous since host factors appear to be of greater importance.14

A system which is weighted towards physiological status would appear to be of potential benefit in assessing the outcome after orthopaedic surgery. Such a system would allow comparison based on the patient’s physiological status and an assessment of the magnitude of surgery and its timing.

Between 1996 and 1998 we used methods similar to those which we have described previously in the development of the POSSUM system,1 in order to design a severity score which would allow the logistic regression equation used in general surgery to be applied to orthopaedics. We assessed 22 surgical severity factors which were later reduced by multivariate analysis to the minimum number necessary to produce an accurate estimate of mortality and morbidity. The resulting operative severity score was similar to the system for general surgery, although individual factors and weightings differed. In this study we have now attempted to validate the application of this new method.

Patients and Methods

Every patient admitted to Warrington Hospital over a period of 12 months on whom orthopaedic surgery had been performed, was assessed using the new orthopaedic POSSUM system. Day-care patients or those treated on an outpatient basis were excluded.

The system includes a physiological assessment and an analysis of operative severity (Table I). The former has 12 variables each of which is divided into four grades with an exponentially increasing score value (1, 2, 4 and 8). Almost all the score variables were available for every
patient, but when a figure was missing, a score of 1 was allocated. The operative severity score includes six variables, each divided into four grades, with an exponentially increasing value (Table I). Definitions of operative complexity are given in Table II. When an operation was not listed the most nearly similar group was chosen. The number of operative variables occurring within 30 days of the surgery demonstrates the chronology of procedures. The operative severity score was available for all patients, although histological confirmation was required to complete the scores for some.

The outcome was assessed as 30-day morbidity and mortality, which allowed comparability with the system for general surgery and could be checked against the 30-day mortality statistics of the Department of Health. The presence of the following complications were recorded as morbidity.

1) Infection: chest, urinary, wound, joint, bony, septicemia or pyrexia of unknown origin.
2) Haemorrhage: deep or superficial.
3) Other wound problems: seroma requiring drainage, dehiscence.
4) Thromboembolic complications: deep-venous thrombosis, pulmonary embolus, cerebrovascular accident.
5) Cardiac: failure, myocardial infarction, arrhythmia.
6) Respiratory: failure, atelectasis, pneumonia.
7) Renal: failure, retention of urine.
8) Unanticipated displacement of an implant.

Exact definitions have been described previously. Predictions of mortality and morbidity for individual patients were estimated using the following equations, in which R1 relates to mortality and R2, morbidity.

\[
\log_e R1/(1-R1) = -7.04 + \left(0.13 \times \frac{\text{physiological score}}{100}\right) + \left(0.16 \times \frac{\text{operative severity score}}{100}\right)
\]

\[
\log_e R2/(1-R2) = -5.91 + \left(0.16 \times \frac{\text{physiological score}}{100}\right) + \left(0.19 \times \frac{\text{operative severity score}}{100}\right)
\]

### Results

Over a period of 12 months there were 2326 orthopaedic operations of which 44% were elective and 56% urgent or emergency procedures. The overall rate of mortality was 2.2%, of which 0.2% were in the elective group and 3.8% in the urgent or emergency group. The overall rate of morbidity was 10.8%, 4.4% for elective and 15.9% for urgent or emergency cases. Table III gives the profile of complications.

The POSSUM logistic regression equations yield an overall predicted mortality of 53 patients versus 51 observed and a predicted morbidity in 254 patients versus 252 observed. The risk spectra for both mortality and morbidity are shown in Table IV, and follow the expected exponential distribution.

The number of operations performed by each of the six surgeons contributing to the study is shown in Table V. As can be seen there are differences in each surgeon’s overall.

---

### Table I. Physiological and operative severity assessment in the orthopaedic POSSUM system

<table>
<thead>
<tr>
<th>Age (yrs)</th>
<th>&lt;60</th>
<th>61 to 70</th>
<th>&gt;71</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac signs</td>
<td>Normal</td>
<td>On cardiac drugs or steroid</td>
<td>Oedema</td>
<td>Warfarin</td>
<td>Raised JVP*</td>
</tr>
<tr>
<td>Chest radiograph</td>
<td>Normal</td>
<td>Borderline cardiomegaly</td>
<td>Cardiomegaly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resp signs</td>
<td>Normal</td>
<td>SOB† exertion</td>
<td>SOB stairs</td>
<td>SOB rest</td>
<td></td>
</tr>
<tr>
<td>Systolic BP (mmHg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulse (min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coma score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood urea (mmol/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood Na (mmol/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood K (mmol/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hb (g/100ml)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White cell count (&gt;1012/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECG</td>
<td>Normal</td>
<td>AF§ (60 to 90)</td>
<td>Any other change</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Jugular venous pressure
†Shortness of breath
‡Chronic obstructive airways disease
§Atrial fibrillation
rates for mortality and morbidity. When corrected for case mix using the POSSUM system, however, there is little difference between the observed rates and those predicted by POSSUM (Table VI).

The predictive accuracy of these equations was assessed by determining the receiver-operating-characteristic curves (ROC curves) and the classification matrices for different levels of predicted mortality and morbidity. The resultant ROC curves are illustrated in Figures 1 and 2, which show good correlation across the range.

**Discussion**

Worldwide public and political interest has been turned on the assessment of quality of care and surgical outcome. This is perhaps easier within the surgical specialties since death after surgery is an obvious adverse outcome. This has led many non-surgical clinicians to suggest that rates of mortality are a suitable indicator of surgical prowess. There is, however, increasing awareness that ‘raw’ data may be both erroneous and deceptive, while morbidity is often ignored. Death after orthopaedic surgery is rare and usually follows a number of antecedent complications.

The overall rates of mortality and morbidity for our period of study appear to be in keeping with previously published rates for inpatient surgery in district general hospitals \(^1,^6\) and the risk profile shows similarities to those...
of general surgery, suggesting that our study group was comparable with the average orthopaedic case mix for a district general hospital. In general surgery, however, there is a tendency to have a greater number of patients whose risk of death is greater than 70%.

Our study shows a close correlation between the overall observed rates for mortality and morbidity and the predictions derived from the POSSUM logistic regression equations. The physiological variables assessed were those shown by logistic regression analysis to be the most important in predicting mortality and morbidity. Other additional variables were not found independently to improve the predictive ability of the logistic regression equation. We have not compared the POSSUM predictions with P-POSSUM. It has been shown previously that if the correct mathematical model is applied there are no significant differences in the predictive ability of either method.

There would appear at first sight to be significant differences in outcome between the six surgeons studied. Rates of mortality varied between 1.1% and 3.0% and of morbidity between 4.6% and 13.4%. Similar variations have been shown for general surgery and can be explained on the basis of the case mix and operative complexity. The orthopaedic POSSUM predictions in our analysis suggest that a similar explanation may account for apparently marked differences in surgical outcome in orthopaedic surgery. This is shown by the close correlation between observed and predicted rates for both mortality and morbidity for individual surgeons.

The ROC curves suggest that the orthopaedic POSSUM score is equally applicable across the spectrum of surgical risk and that the predictions for an individual patient may be of benefit when assessing a patient who has died or suffered a complication. We have previously drawn attention to the benefits of audit in patients who survive, but for whom a predicted risk of death exceeds 50%, since data on these patients are often of more value than those on patients who die. Recently, Wilson et al have shown that potentially significant improvements in overall care are possible by optimisation in this particular group.

Clearly, no regression equation for risk assessment should remain static indefinitely. We have found no need for change over the past ten years, but should dramatic changes occur in the future the equation can be easily updated without the need to alter the score variables. Should this occur, patients already scored would be assessed with the present equation and new patients with an updated format.

Our study demonstrates that POSSUM can be used as an audit aid to assess the quality of orthopaedic care. Thus a quality measurement similar to that used in general surgery, namely the ratio of observed adverse events to predicted adverse events, may be used in orthopaedic surgery and be more sensitive than simple rates of mortality and morbidity.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References