

# Comparative audit: the trouble with POSSUM

Simon D Bann BSc FRCS Sanjeev Sarin MS FRCS

*J R Soc Med* 2001;94:632-634

## SUMMARY

The Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (POSSUM) has been proposed for use in comparative audit between surgeons and between hospitals. To assess its feasibility, POSSUM scoring was attempted on admission in all patients under the care of two consultant surgeons over a six-month period. Scores were awarded only if all investigations necessary for POSSUM were performed; investigations unnecessary for effective treatment were not performed.

815 patient discharges were recorded over the six-month period, with 521 patients undergoing operative procedures. Of those undergoing an operation, scores could be allocated in only 155 (30%).

Scoring systems such as POSSUM are procedure-based, thereby excluding those who do not undergo an operation. However, most of our operative cases were also excluded. Full POSSUM scoring will often require additional investigations. POSSUM is unlikely to be of use in the wider setting of comparative audit.

## INTRODUCTION

Scoring systems in all areas of medicine are receiving close attention because of the need to evaluate and monitor healthcare delivery and outcomes. The main application is in comparative audit, but an effective scoring system could be useful for other purposes—for example in research, to standardize for case mix, and in clinical management as a prognostic indicator.

In surgery various risk scores are used but a recent review article recommended the Physiological and Operative Severity Score for the enUmeration of Morbidity and Mortality (POSSUM) as the most appropriate for general surgical practice<sup>1</sup>. POSSUM was originally described by Copeland *et al.* in 1991<sup>2</sup> and its Portsmouth modification in 1996<sup>3,4</sup>. Both systems take the same 12 physiological variables and combine these with 6 operative variables (Box 1). Many of the evaluations have been in surgical subspecialties; and, provided the data are correctly analysed, POSSUM has been found accurate in predicting outcomes<sup>5-8</sup>. Midwinter *et al.*<sup>8</sup> comment: 'Such equations could also be used by an individual surgeon or unit to assess his/her performance'. We have examined the suitability of POSSUM for use in comparative audit within a general surgical setting.

Box 1 POSSUM physiological and operative indices

Physiological	Operative
Age	Operative severity
Cardiac history	Multiple procedures
Respiratory history	Total blood loss
Blood pressure	Peritoneal soiling
Pulse rate	Presence of malignancy
Glasgow coma score	Mode of surgery
Haemoglobin	
White cell count	
Urea	
Sodium	
Potassium	
Electrocardiogram	

## PATIENTS AND METHODS

Over six months, all patients admitted under the care of two general surgical consultants were included. On admission they were scored by POSSUM criteria on a datasheet containing all the physiological variables. Patients transferred to the care of the firm whilst an inpatient and patients whose care was shared with another firm were included. Upon discharge the datasheet was completed by assignment of the operative scores, if appropriate. At weekly meetings of the entire two surgical firms, each patient discharged during the previous week was discussed and any inaccuracies were corrected. Completed datasheets and discharge summaries were held on a computerized database (SIS Plus, London), which was updated weekly. The hospital patient administration system (PAS) was not relied upon.

Department of Surgery, Watford General Hospital, Vicarage Road, Watford WD1 8HB, UK

Correspondence to: Mr S Sarin  
E-mail: SSarin@compuserve.com

Emergency admissions were investigated as thought appropriate by the admitting team. Preoperative routine investigation policy for elective admissions in our unit is as follows:

1. All patients (other than those having minor operations under local anaesthesia) to have a haemoglobin estimation.
2. Urea and electrolytes measured if clinically indicated—e.g. fluid balance problems or taking diuretics or history of renal disease.
3. Electrocardiograms on all patients over 50 years of age (and under that age if clinically indicated).
4. Chest X-ray only if a clinical indication—e.g. suspected pneumonia.

These guidelines are in keeping with those of most surgical departments and we did not deviate from our routine practice. We included all surgical cases, so no part of the workload was unaccounted for. Any patient who did not undergo the necessary investigations could not be given a physiological score.

**RESULTS**

Over the six months, 815 patients were discharged with 27 deaths (3.3%). In total there were 390 emergency admissions and 398 elective admissions and 27 patients were internal or external transfers (see Table 1). Not all patients had an operative procedure: some merely underwent investigation or non-operative treatment. This group included all those undergoing radiological investigation and intervention since they would not receive a POSSUM score by definition—though they constitute a substantial part of the surgical workload.

521 operations were performed on 501 patients, 162 of these classified as emergency and 342 as elective. The remaining 17 operations were performed on transferred patients. The American Society of Anesthesiologists (ASA)

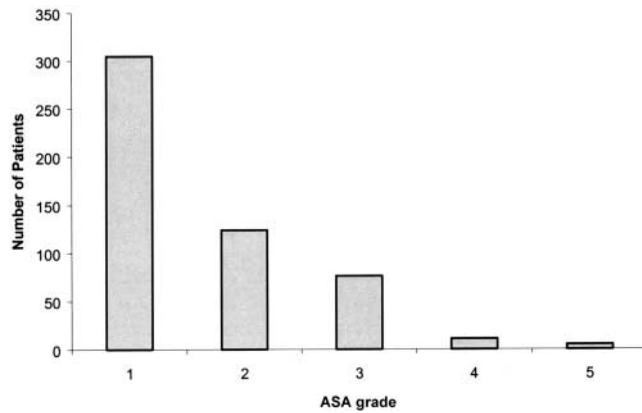


Figure 1 Breakdown of patients by ASA grade

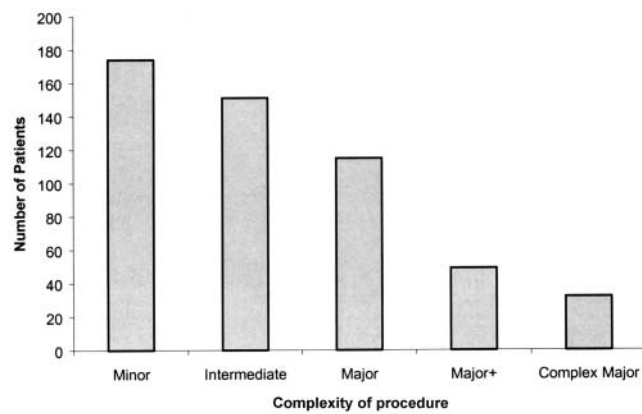


Figure 2 Breakdown of patients by complexity of procedure (BUPA operative severity score)

score for these patients is shown in Figure 1 and the level of complexity of the operations in Figure 2. A complete POSSUM score was obtained in 155 patients (29.8%). Emergency admissions accounted for 72 of these and elective admissions for 68. The remainder, 15, were from the transfer group.

**DISCUSSION**

This prospective evaluation casts serious doubt on the suitability of POSSUM for use in audit of general surgeons or their units. The reasons for the disappointing performance seem to lie in the history of the scoring system. Firstly, in formulating the original scoring system, Copeland *et al.* excluded all patients who were admitted for less than 24 hours (unknown number) as well as those having trauma surgery. Thus day cases would have been excluded, as they were specifically in P-POSSUM along with children<sup>3,4</sup>. Day cases and children make up 48% of our workload. Furthermore, in Copeland's study all the patients had blood samples taken for determination of urea and electrolytes, haemoglobin concentration and white cell count, and all had electrocardiography performed. This

Table 1 Breakdown of surgical admissions

Admissions	No.
<i>Elective</i>	
Adult inpatient	184
Day case	199
Paediatric	15
Total	398 (48.8%)
<i>Emergency</i>	
Adult	359
Paediatric	31
Total	390 (47.9%)
<i>Transfers</i>	27 (3.3%)
<b>Total</b>	<b>815</b>

blanket approach to preoperative investigation for inpatient treatment is not in keeping with our hospital guidelines—adding to workload with little benefit in terms of diagnosis or operative risk (for example, an electrocardiogram in a fit young adult undergoing appendicectomy). Further exclusion of those inpatients in our study who underwent operative treatment but did not have the full POSSUM work-up left a net scored population of about 30%. To improve the scoring rate, some have suggested that a test deemed unnecessary could be assumed normal and allocated a score of one, but the original reports imply that all tests necessary for a POSSUM score were performed<sup>2-4</sup>. In keeping with these reports, we chose not to allocate a 'minimum' score to those patients with an incomplete POSSUM physiological dataset.

Whilst it is true that our patients who were not scored probably represent those who underwent minor or intermediate surgery, if outcome is measured by any event other than mortality (e.g. complications) exclusion of such a large group would seriously affect analysis of the results. Also, deaths in the low-risk groups may flag important differences in care. For any comparative audit system to be effective, differences in morbidity and mortality should be especially highlighted in this group. Even if this group were included by assigning normal values for the physiological variables, both POSSUM and P-POSSUM would exaggerate the expected mortality—since the minimum risk of death with POSSUM is 1.08% and with P-POSSUM 0.20%. This low-risk group makes up the bulk of current general surgical practice, so surgeons would constantly be performing better than predicted.

Another factor is timing. For POSSUM, Copeland *et al.*<sup>2</sup> scored their patients immediately before surgery, allowing for resuscitation post admission, whereas for P-POSSUM Whiteley *et al.* scored on admission<sup>3,4</sup>. Recent snapshot surveys based on the POSSUM dataset have been unclear about the time of scoring, and some groups have erroneously applied the P-POSSUM equation (derived from data collected on admission) to data scored preoperatively. Since preoperative intervention affects outcome<sup>9,10</sup>, ambiguity of this kind will distort the results. Even if the time of scoring were standardized to be preoperative, a surgical firm that aggressively pre-optimized patients would not be shown to be providing better care, only to be operating on fitter patients than the average.

Finally, how does POSSUM perform in demonstrating the competence or otherwise of the surgical team? The score allocates additional points if the surgeon spills more blood and operates on the patient many times (perhaps because of complications)<sup>5</sup>. Copeland in his original paper assigns an operative score at discharge but does not clarify how this score is achieved: is it the score on opening the

abdomen, and if there are multiple procedures does one enter the highest score for the factors or the one at the last operation? Supposing, for example, a man of 65 requires a left hemicolectomy for colonic obstruction. His initial physiological score is 15. Surgeon A operates with little blood loss and no peritoneal soiling—operative score 13. Surgeon B does the same procedure with greater blood loss and with peritoneal soiling, and the patient has to return to theatre with a colonic leak—operative score 30. Inserting these scores into the POSSUM equation gives an expected mortality of 5% for surgeon A and 43% for surgeon B. While POSSUM may correctly predict the expected risk of death in each scenario it is of little value in flagging up surgical failings: surgeon B appears to be operating on more technically challenging patients. Put another way, what the POSSUM dataset appears to ask is 'Given this *patient* and *surgeon*, what is the expected outcome?' The question that any effective system for comparative audit needs to ask is 'Given this *patient*, what is the expected outcome?' In conclusion, while the POSSUM and P-POSSUM equations are potentially accurate in predicting mortality after an operation, in our study a large proportion of surgical patients were excluded by application of our standard criteria for investigation. Use of the POSSUM dataset effectively downplays surgeon-related variables. In our surgical practice it is not suitable for comparative audit.

## REFERENCES

- 1 Jones HJS, de Cossart L. Risk scoring in surgical patients. *Br J Surg* 1999;**86**:1499–57
- 2 Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg* 1991;**78**:355–60
- 3 Whiteley MS, Prytherch DR, Higgins B, Weaver PC, Prout WG. An evaluation of the POSSUM surgical scoring system. *Br J Surg* 1996;**83**: 812–15
- 4 Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG and Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. *Br J Surg* 1998;**85**:1217–20
- 5 Wijesinghe LD, Mahmood T, Scott DJA, Berridge DC, Kent PJ, Kester RC. Comparison of POSSUM and the Portsmouth predictor equation for predicting death following vascular surgery. *Br J Surg* 1998;**85**:209–12
- 6 Copeland GP, Jones D, Wilcox A, Harris PL. Comparative vascular audit using the POSSUM scoring system. *Ann R Coll Surg Engl* 1993; **75**:175–7
- 7 Sagar PM, Hartley MN, Mancey-Jones B, Sedman PC, May J, Macfie J. Comparative audit of colorectal resection with the POSSUM scoring system. *Br J Surg* 1994;**81**:1492–4
- 8 Midwinter MJ, Tytherleigh M, Ashley S. Estimation of mortality and morbidity in vascular surgery using POSSUM and the Portsmouth predictor equation. *Br J Surg* 1999;**86**:471–4
- 9 McLroy B, Miller A, Copeland GP, Kiff R. Audit of emergency preoperative resuscitation. *Br J Surg* 1994;**81**:200–2
- 10 Wilson J, Woods I, Fawcett J, *et al.* Reducing the risk of major elective surgery: randomised controlled trial of preoperative optimisation of oxygen delivery. *BMJ* 1999;**318**:1099–1103