

## Decisions to operate: the ASA grade 5 dilemma

**J Horwood, S Ratnam, A Maw**

Department of General Surgery, Glan Clwyd Hospital, Rhyl, UK

### ABSTRACT

**INTRODUCTION** Deciding to operate on high risk patients suffering catastrophic surgical emergencies can be problematic. Patients are frequently classed as American Society of Anesthesiologists (ASA) grade 5 and, as a result, aggressive but potentially lifesaving intervention is withheld. The aim of our study was to review the short-term outcomes in patients who were classed as ASA grade 5 but subsequently underwent surgery despite this and to compare the ASA scoring model to other predictors of surgical outcome.

**METHODS** All patients undergoing emergency surgery with an ASA grade of 5 were identified. Patient demographics, indications for surgery, intraoperative findings and outcomes were recorded. In addition to the ASA scores, retrospective Portsmouth Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (P POSSUM) and Acute Physiology and Chronic Health Evaluation II (APACHE II) scores were calculated and compared to the observed outcomes.

**RESULTS** Nine patients (39%) survived to discharge. ASA grade was a poor predictor of outcome. P POSSUM and APACHE II scores correlated significantly with each other and with observed outcomes when predicting surgical mortality. The median stay for survivors in the intensive care unit was nine days.

**CONCLUSIONS** In times of an ageing population, the number of patients suffering catastrophic surgical events will increase. Intervention, with little hope of a cure, a return to independent living or an acceptable quality of life, leads to unnecessary end-of-life suffering for patients and their relatives, and consumes sparse resources. The accuracy and reliability of ASA grade 5 as an outcome predictor has been questioned. P POSSUM and APACHE II scoring systems are significantly better predictors of outcome and should be used more frequently to aid surgical decision-making in high risk patients.

### KEYWORDS

ASA 5 – Outcome prediction scores – P POSSUM

Accepted 22 March 2011

### CORRESPONDENCE TO

**James Horwood**, Specialist Registrar, Department of General Surgery, Glan Clwyd Hospital, Rhyl, Denbighshire LL18 5UJ, UK  
E: tipper111@aol.com

Deciding to operate on high risk patients with various surgical emergencies can be problematic. In times of an ever ageing population, the numbers of patients with multiple medical co-morbidities presenting to the general surgical 'take' having suffered a catastrophic surgical event, are likely to increase.<sup>1,2</sup> Surgical intervention, with little hope of a cure, a return to independent living or an acceptable quality of life, may lead to unnecessary end-of-life suffering for the patients and their relatives, and consumes significant financial resources.

Many practitioners have developed preoperative prediction models aimed at discriminating those patients most likely to survive surgical intervention and return to their pre-morbid level of function from those in whom aggressive, invasive treatment is likely to be futile.<sup>3,4</sup> Since its introduction,<sup>5</sup> the American Society of Anesthesiologists (ASA) scoring system has been used ubiquitously for preoperative risk stratification and remains the most common system in everyday use. Despite this, many authors have questioned the accuracy and inter-observer reliability of the ASA system.<sup>6–8</sup> Nevertheless, patients continue to be classed as ASA grade 5 and, as a result, aggressive but potentially lifesaving

intervention is withheld due to the presumed futility of the attempt.

The aim of our study was to review short-term outcomes in patients from our institution who, despite being classed as ASA grade 5, underwent emergency surgery and to compare the ASA scoring model with other predictors of outcome.

### Methods

Between October 2004 and November 2007 the hospital operations database was searched for all patients given a preoperative ASA score of 5 and the notes of these patients were reviewed. Patient demographics (age, sex, body mass index), presenting diagnosis, intraoperative findings and procedures, length of stay on the intensive care unit (ICU), hospital stay, and dates and causes of death were recorded.

In addition to the ASA score, retrospective Acute Physiology and Chronic Health Evaluation II (APACHE II),<sup>9</sup> Portsmouth Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (P POSSUM)<sup>5</sup> and Hardmam<sup>10</sup> scores were calculated. These scoring systems were chosen following an extensive literature review that

Table 1 P POSSUM sensitivity, specificity and correct classification

Scoring cut off	Sensitivity	Specificity	Correctly classified
4%	100%	0%	60.9%
20%	100%	22.2%	65%
50%	100%	77.8%	91.3%
65%	78.5%	88.9%	82.6%
85%	35.7%	100%	60.9%

suggested that these risk stratification models were the most widely used and studied. Statistical analysis using SPSS® v15 (SPSS Inc, Chicago, Illinois, US) was used to compare the accuracy of the various scoring systems in predicting the observed outcomes.

## Results

Between October 2004 and November 2007, 23 patients with a perioperative ASA score of 5 underwent emergency surgery. The median age was 76 years (range: 64–93 years), the male to female ratio was 2:1 and the median body mass index was 33kg/m<sup>2</sup>.

ASA scores were allocated by the most senior anaesthetist in attendance. In all but two cases, this was an experienced consultant. The decision to operate was made by the consultant surgeon and anaesthetist with responsibility for the patient's care, following discussion with the patient and/or their relatives and taking their wishes into consideration.

The intraoperative diagnosis, ICU admission, outcome prediction model scores and observed outcomes were recorded. (See online appendix.) Four patients (17%) were found to have suffered a ruptured abdominal aortic aneurysm (AAA). Of these, one patient (25%) survived to discharge. The three patients who died had Hardman index scores of 75% for predicted mortality. The patient who survived had a Hardman predicted mortality of 25%. However, P POSSUM and ASA scores failed to discriminate the survivor (P POSSUM predicted mortality of 82%).

Three patients (13%) had a sigmoid colonic perforation secondary to diverticular disease. One patient survived to discharge, with the P POSSUM and APACHE II systems predicting 18% and 17% mortality respectively. Two fatalities were due to multiorgan failure with P POSSUM scores of 93% and 50% and APACHE II scores of 85% and 64% for predicted mortality respectively.

Two patients (9%) suffered an anastomotic leak after an elective bowel resection. Both patients were assigned an ASA score of 3 at the time of their original operation. One patient survived to discharge, with an APACHE II predicted mortality score of 20% and a P POSSUM predicted mortality of 15%.

Three patients (13%) were found to have mesenteric ischaemia. One, who survived to discharge, had P POSSUM and APACHE II scores predicting 4% and 8% for mortality respectively.

Another patient (case 18), with a predicted mortality of 60% and 72% on the P POSSUM and APACHE II systems

respectively, was taken to theatre for drainage of a large abdominal wall abscess but the collection was found to be communicating with the abdominal cavity. Due to the patient's age and multiple co-morbidities, the decision was taken by the operating surgeon not to explore the peritoneum. The patient died of sepsis and multiorgan failure within 12 hours of surgery.

An additional patient, who had undergone a laparoscopic radical prostatectomy (ASA grade 3), developed signs of hypovolaemic shock that, at laparotomy, was found to be secondary to a bleeding port site. He survived to discharge. He was assigned an ASA score of 5. However, P POSSUM and APACHE II scores of 45% and 28% for predicted mortality were recorded.

Admitted with peritonitis, a further patient (case 23) had previously undergone a left hemicolectomy for a T4 adenocarcinoma of the colon. Following discussion with the patient and her relatives, a laparotomy was performed and she was found to have widespread intra-abdominal recurrence of carcinoma. She survived to discharge from hospital but with a poor long-term prognosis.

Nine patients (39%) given a preoperative ASA score of 5 survived to discharge from hospital, suggesting that ASA was a poor predictor of clinical outcome.

Receiver operator characteristic (ROC) curves were constructed to compare the P POSSUM and APACHE II models for surgical outcome prediction in our patient cohort (Fig 1). The area under the P POSSUM curve was 0.91 (95% confidence interval = 0.78–1.00) and for APACHE II it was 0.87 (95% confidence interval = 0.72–1.00). The difference between P POSSUM and APACHE II failed to reach statistical significance ( $p=0.65$ ).

Table 1 demonstrates the sensitivity, specificity and correct classification (ie survival/mortality) at various levels of outcome prediction for the P POSSUM model. This shows that at a P POSSUM score of 50% predicted mortality, 91% of patients were correctly classified (ie correct prediction of mortality). Similar tests for APACHE II (Table 2) demonstrated that at a score of 45% predicted mortality, 86% of patients were correctly classified.

This suggests that, excluding ruptured aneurysms, had we disregarded the ASA scoring system and only offered surgery to patients with a P POSSUM score indicating a predicted mortality of less than 50%, we would have operated on every patient who eventually survived to discharge with the exception of one patient and denied operations to every patient who eventually died.

Table 2 APACHE II sensitivity, specificity and correct classification

Scoring cut off	Sensitivity	Specificity	Correctly classified
8%	100%	0%	60.9%
26%	92.8%	33.3%	69.6%
45%	85.7%	88.9%	87.0%
62%	35.7%	100%	60.9%
72%	21.4%	100%	52.2%

## Discussion

The population of elderly people in the UK is increasing.<sup>11</sup> This has resource and cost implications for the National Health Service. It is therefore reasonable to expect the numbers of elderly patients, with multiple medical co-morbidities, presenting on the general surgical emergency 'take' to increase considerably.

The decision on whether to operate on sick elderly patients with an intra-abdominal emergency is one of the most difficult in general surgery. Needless performing major abdominal surgery in patients with little hope of a cure or a return to their pre-morbid function causes significant end-of-life suffering for patients and considerable distress to their relatives. In addition, precious theatre time and costly critical care beds occupied by such cases may deny these facilities to patients with a better prognosis. However, clinicians would not wish to decline potentially lifesaving surgery to patients who have at least a small chance of survival.

Many confounding factors influence the way in which the general population perceives risks from operative intervention.<sup>12,15</sup> Perception of risk can be influenced by a multitude of factors that do not necessarily impact on the outcome itself. The National Confidential Enquiry into Perioperative Deaths demonstrated the poor assessment of risk, finding that surgeons perceived an increase in perioperative risk in only 66% of patients who eventually died.<sup>14</sup> Accurate risk assessment in general surgical patients undoubtedly aids the decision making process, allowing a balance of the risks of surgery against the potential benefits to be considered objectively by the practitioner, patient and relatives in a collective decision making process.

Many surgical predictors of morbidity and mortality have been developed. The most universally used predictor remains the ASA scoring system.<sup>5</sup> Many authors have demonstrated that the ASA system is robust, simple to apply and predictive of surgical outcome.<sup>15,16</sup> However, the subjective nature of the system and lack of clarity between the various ASA grades (particularly grades 2 and 5) has led to considerable inter-rater disagreement. In addition, no provision is made for patients suffering from more than one chronic medical condition and the ASA classification does not consider the acute pathology that has precipitated the emergency admission.

Patients classed as ASA grade 5 are not expected to survive 24 hours, with or without surgery. It would therefore be a reasonable approach to withhold surgery from this highly selected group. However, as our study has demonstrated,

40% of patients classed as ASA grade 5 who did undergo surgery survived to discharge. Therefore, in this emergency surgery patient group, the ASA system was a poor predictor of clinical outcome.

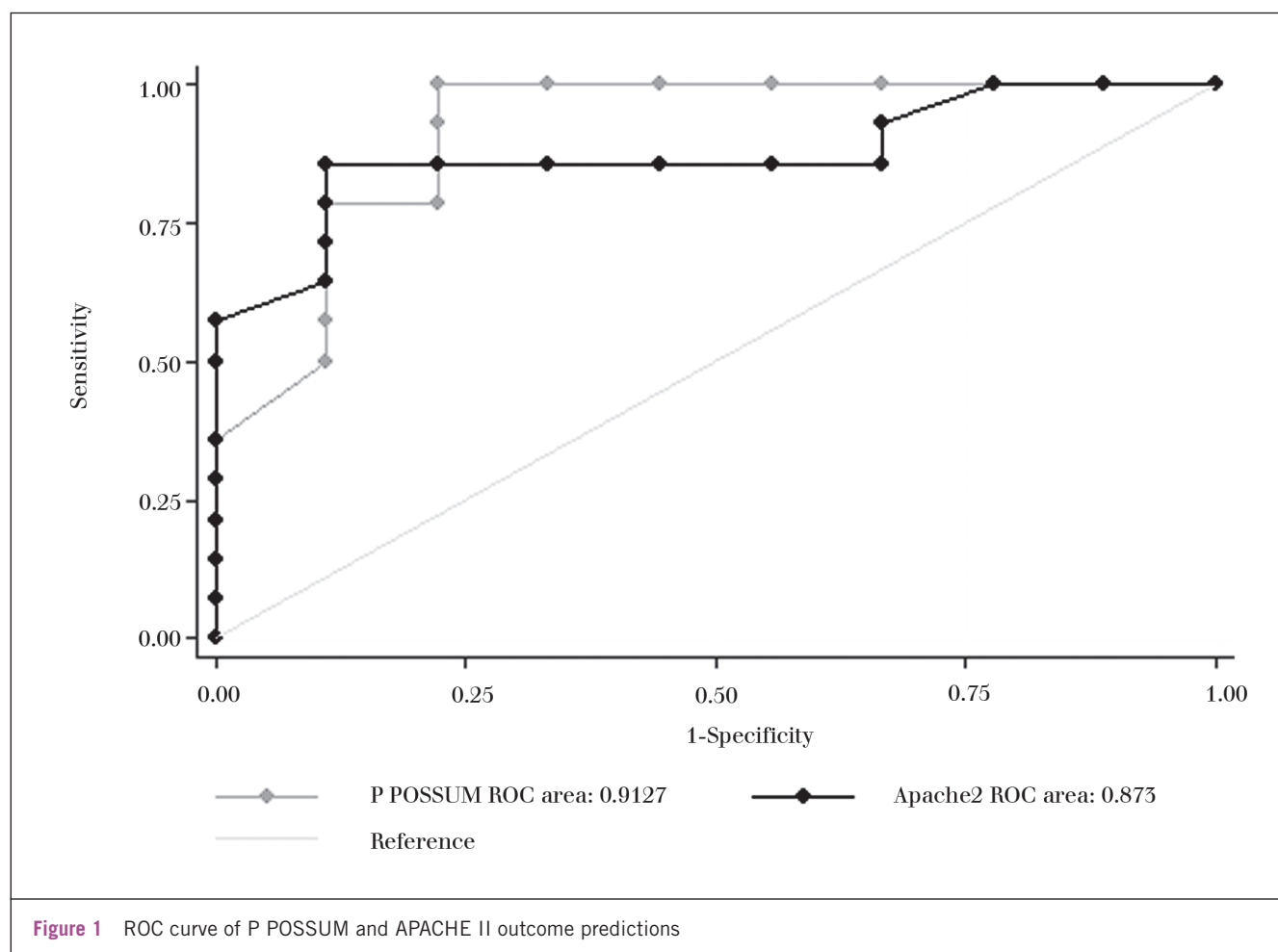
Poor inter-observer reliability with the ASA system was observed in our study. Cases 15 and 21, suffering major complications following elective surgery, had their ASA scores increased from 2/3 to 5 for emergency reoperation. On informal review of these cases, some experienced clinicians felt that the original ASA scores should have been kept but with the E (emergency) suffix applied. Such inter-rater disagreement could result in considerable confusion if the surgeons were following a non-operative policy in all patients classed as ASA grade 5.

The ASA score is subjective in nature and does not take into account all the pathological or physiological processes occurring, which have a clear impact on outcome. Case 5, a relatively young and previously fit patient with perforated sigmoid diverticular disease with non-faecal peritoneal soiling, was assigned ASA grade 5. However, both the P POSSUM and APACHE II scores identified that this patient had an acceptable chance of survival (82% and 83% respectively).

The P POSSUM score has been shown consistently to be an accurate predictor of surgical outcome,<sup>17</sup> taking into account not only the pre-morbid status of the patient but also physiological parameters pertaining to his or her current admission and gross intraoperative findings that influence outcome (eg faecal peritonitis).

The P POSSUM score has undoubtedly aided surgeons in the decision making process, particularly in the high risk and emergency surgery groups. However, as in the case of ruptured AAAs, pathology-specific scoring systems have been developed that are superior to the P POSSUM score for predicting outcome. In our institution, patients presenting with ruptured AAAs are assessed using the Hardman index, which proved superior to the P POSSUM system in predicting operative mortality. Both the Hardman index and Glasgow Aneurysm Score have been shown to accurately predict mortality and have had an impact on the operative decision making for patients with ruptured aneurysms.<sup>18</sup>

The P POSSUM score takes into account the operative findings, which of course are not available to the assessing physician prior to surgical intervention. Therefore, using the P POSSUM score for preoperative decision making requires the surgeon to 'best guess' the operative findings to input into the model. Perhaps modification and validation of a 'preoperative P POSSUM' score, taking into account the



results of investigations (eg computed tomography) could be developed to estimate the operative risk and inform decision making. A prospective study of such a preoperative P POSSUM scoring model would be useful.

Excluding the patients with ruptured aortic aneurysms, taking a P POSSUM score of 50% predicted mortality as a cut off point for intervention correctly classified the outcome in 91% of cases.

The APACHE II score is a validated disease severity classification system whose use is widespread to predict morbidity and mortality. APACHE II is a point score that relies on 12 routine physiologic measurements, age and previous health status to give a score (0–71), predictive of clinical outcome. Many authors have demonstrated the predictive value of the APACHE II system in both general and emergency surgical patients.<sup>19–22</sup> The APACHE II system, however, is rather cumbersome, requiring a considerable number of investigations and patient history, some of which would not be immediately available in the emergency setting. In addition, the APACHE II system fails to take into account operative findings. Although we found that the APACHE II score correlated well with clinical outcome in our high risk population, we would concur with other authors that the

APACHE II system is most practically utilised in the early postoperative period in the ICU and high dependency unit.

The P POSSUM and APACHE II scores correlated well both with observed outcomes and each other (areas under ROC curve = 0.91 and 0.87 respectively). However, in some cases we found considerable variation in the predicted mortality between the two scoring systems. A patient with perforated diverticular disease with free intraperitoneal bowel contents (case 6) scored 95% predicted mortality with the P POSSUM system but only 50% with the APACHE II system. On closer examination of this case, the high P POSSUM score was accounted for by the weighting in the system that is given to intraoperative findings.

One weakness of our study is the retrospective nature of the data collection. When calculating both the P POSSUM and APACHE II scores, we did so with the knowledge of the intraoperative findings, allowing us to obtain a more accurate score. Clearly, however, in the emergency situation when faced with a clinical dilemma, the intraoperative findings and procedure required will not be known and can only be estimated. This may impact on the P POSSUM score achieved preoperatively and is an area deserving of further study.

## Conclusions

In the era of ageing populations, patients with multiple comorbidities having suffered an abdominal catastrophe (and therefore presenting the attending surgeon with difficult management decisions) are likely to increase. Evidence-based surgical risk scoring systems can be a valuable tool to aid the surgeon in making these critical decisions. The ideal scoring system would be simple to apply in the emergency setting, robust, reproducible and, most importantly, it would correlate highly with clinical outcomes.

We found that the ASA system, ubiquitously used for risk prediction, can be unreliable when assessing outcome in our selected group of high risk patients. Despite shortcomings, we found that the P POSSUM system is a practical, valuable and accurate method of risk prediction. We would recommend caution in withholding lifesaving treatment from patients based on an ASA score of 5 and would advocate using risk prediction models for further evidence-based assessment of the risks of surgery. Nevertheless, scoring systems only aid the surgeon in the decision making process and should not replace experience and clinical judgement.

## Appendix

An appendix detailing diagnosis, outcome prediction model scores and observed outcomes is available online.

## References

- Williams JH, Collin J. Surgical care of patients over eighty: a predictable crisis at hand. *Br J Surg* 1988; **75**: 371–373.
- Rørbaek-Madsen M, Dupont G, Kristensen K *et al*. General surgery in patients aged 80 years and older. *Br J Surg* 1992; **79**: 1,216–1,218.
- Whiteley MS, Prytherch DR, Higgins B, Weaver PC, Prout WG. An evaluation of the POSSUM surgical scoring system. *Br J Surg* 1996; **83**: 812–815.
- Knaus WA, Zimmerman JE, Wagner DP *et al*. APACHE – acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med* 1981; **9**: 591–597.
- Saklad M. Grading of patients for surgical procedures. *Anesthesiology* 1941; **2**: 281–284.
- Mak PH, Campbell RC, Irwin MG. The ASA Physical Status Classification: inter-observer consistency. *Anaesth Intensive Care* 2002; **30**: 633–640.
- Aronson WL, McAuliffe MS, Miller K. Variability in the American Society of Anesthesiologists Physical Status Classification Scale. *AANA J* 2003; **71**: 265–274.
- Haynes SR, Lawler PG. An assessment of the consistency of ASA physical status classification allocation. *Anaesthesia* 1995; **50**: 195–199.
- Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985; **13**: 818–829.
- Hardman DT, Fisher CM, Patel MI *et al*. Ruptured abdominal aortic aneurysm: who should be offered surgery? *J Vasc Surg* 1996; **23**: 123–129.
- Population: by age. Office for National Statistics. [www.statistics.gov.uk/ccin/nugget.asp?id=949](http://www.statistics.gov.uk/ccin/nugget.asp?id=949) (cited May 2011).
- Lloyd AJ. The extent of patients' understanding of the risk of treatments. *Qual Health Care* 2001; **10 Suppl 1**: i14–18.
- Lloyd A, Hayes P, Bell PR, Naylor AR. The role of risk and benefit perception in informed consent for surgery. *Med Decis Making* 2001; **21**: 141–149.
- Campling EA, Devlin HB, Hoile RW, Lunn JN. *The Report of the National Confidential Enquiry into Perioperative Deaths 1991/1992*. London: National Confidential Enquiry into Perioperative Deaths; 1993.
- Mella J, Biffin A, Radcliffe AG *et al*. Population-based audit of colorectal cancer management in two UK health regions. *Br J Surg* 1997; **84**: 1,731–1,736.
- Arenal JJ, Bengoechea-Beeby M. Mortality associated with emergency abdominal surgery in the elderly. *Can J Surg* 2003; **46**: 111–116.
- Jones HJ, de Cossart L. Risk scoring in surgical patients. *Br J Surg* 1999; **86**: 149–157.
- Sharif MA, Lee B, Makar RR *et al*. Role of the Hardman index in predicting mortality for open and endovascular repair of ruptured abdominal aortic aneurysm. *J Endovasc Ther* 2007; **14**: 528–535.
- Koperna T, Semmler D, Marian F. Risk stratification in emergency surgical patients: is the APACHE II score a reliable marker of physiological impairment? *Arch Surg* 2001; **136**: 55–59.
- Garcea G, Ganga R, Neal CP *et al*. Preoperative early warning scores can predict in-hospital mortality and critical care admission following emergency surgery. *J Surg Res* 2010; **159**: 729–734.
- de Cássia Braga Ribeiro K, Kowalski LP. APACHE II, POSSUM, and ASA scores and the risk of perioperative complications in patients with oral or oropharyngeal cancer. *Arch Otolaryngol Head Neck Surg* 2003; **129**: 739–745.
- Chandra A, Mangam S, Marzouk D. A review of risk scoring systems utilised in patients undergoing gastrointestinal surgery. *J Gastrointest Surg* 2009; **13**: 1,529–1,538.