ANZ J. Surg. 2008; **78**: 535–539 doi: 10.1111/j.1445-2197.2008.04567.x

SPECIAL ARTICLE

EVALUATION OF SURGICAL PERFORMANCE USING V-POSSUM RISK-ADJUSTED MORTALITY RATES

DAMIEN MOSQUERA, NATHANIEL CHIANG AND ROBERT GIBBERD

Department of Surgery, Taranaki Base Hospital, New Plymouth, New Zealand

Vascular-Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (V-POSSUM) is a risk-adjusted scoring system for predicting 30-day mortality in patients undergoing vascular surgery. It can assess surgical performance by comparing predicted deaths with observed deaths. The aim of this analysis was to assess trends in surgical performance over time using risk-adjusted 30-day mortality as the primary outcome. Major vascular surgery procedures (n = 454) were prospectively scored for V-POSSUM between 1995 and 2006. Procedures were divided into 11 consecutive time bands. Observed and predicted deaths were compared using the logistic regression equation derived for V-POSSUM. The observed death rates decreased over time, as did the predicted number of deaths calculated from the V-POSSUM scores. The overall predicted mortality rate was 17.2% and the rate varied with the 12-month period, with a high of 23.9% and a low of 9.2%. The downward trend in the predicted rate shows that the patient risk factors have changed over time and that the risk of dying has declined by almost 50% (from 21.6 to 11.1%). There was a trend towards improved surgical performance over time, with a drop in the observed to predicted ratios of deaths. Observed and predicted deaths changed over the study periods. There was a trend towards improved performance compared with the risk-adjusted predicted mortality. V-POSSUM is a useful tool in the longitudinal assessment of performance in major vascular surgery.

Key words: mortality, risk adjustment, severity of illness, survival rate, trend.

Abbreviations: OS, operative score; POSSUM, Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity; PS, physiological score.

INTRODUCTION

Monitoring clinical performance is increasingly important and forms an essential part of the Royal Australasian College of Surgeons Continuing Professional Development Programme.[1] One frequently quoted marker of surgical performance is mortality rate after surgery. Unadjusted mortality rates, although relatively easy to obtain, do not take into account the type of surgery, patient selection by the surgeon and patient comorbidity. There is a clear need to adjust for variable casemix when calculating mortality rates, so that realistic comparisons of performance can be made.

The Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (POSSUM) risk-adjusted scoring system for predicting mortality in surgical patients was developed by Copeland *et al.* in 1991. Mortality risk is predicted using a 12-factor, 4-grade, physiological score (PS) and a 6-factor, 4-grade operative score (OS).[2] By adjusting for the severity of the operation and the physiological status of the patient, predicted outcomes can be compared with those observed in practice, providing a more reliable assessment of performance. The original POSSUM mortality regression equation has been modified by the Vascular Surgical Society of Great Britain and Ireland to produce a regression equation (V-POSSUM) that can be specifically applied to patients undergoing major vascular surgery.[3] The system can be used to compare surgical outcomes, particularly 30-day mortality.

Reliable surgeon-specific mortality rates are important for several reasons. They may be important in obtaining informed consent and reassuring patients. Primary health-care referrers, purchasers and health insurers can be satisfied on standards of care that they are requisitioning. Comparing performance at different periods within a career and benchmarking against surgeons in similar specialties can assist individual surgeons as their practice and patient groups alter during their careers. Reliable adjusted mortality figures may support surgeons operating on difficult cases, where they otherwise may be discouraged, if only high crude mortality figures are reported.

Studying and reporting individual surgeon performance measures has become an increasingly important component for clinical governance and this paper looks at these issues by reporting such data for a single surgeon (D. M.). The aim of this study was to assess whether the benefits were realized. Furthermore, to monitor the changes that have occurred in one individual surgeon's performance of vascular surgery after appointment to a specialist post, V-POSSUM risk-adjusted 30-day mortality was used as the primary outcome measure. In particular, we look at nearly 10 years of data and review the results when data are aggregated over time and examine fluctuations in performance.

METHODS

Data were collected prospectively by a single surgeon (D. M.), who was appointed to a specialist vascular surgery post in September 1995 at Birmingham Heart of England National Health Service Foundation Trust in the UK, a large district

D. Mosquera MD, FRCS, FRACS; **N. Chiang** MB ChB; **R. Gibberd** PhD.

Correspondence: Mr Damien Mosquera, Department of Surgery, Taranaki Base Hospital, David Street, New Plymouth 4310, New Zealand.
Email: damien.mosquera@tdhb.org.nz

Accepted for publication 6 December 2007.

general teaching hospital serving a population of approximately 450 000. Initially, emergency vascular cover was shared between three surgeons, but by 2000, five vascular surgeons shared emergency cover. In February 2002, D. M. moved to Taranaki Base Hospital in New Plymouth, New Zealand. Base Hospital is a 250-bed district hospital serving a population of approximately 105 000. This post is a general surgery post with an interest in vascular surgery and an effective vascular emergency roster of one in two.

The major vascular procedures collected date between September 1995 and April 2006 and are an uninterrupted consecutive series. Only procedures where the surgeon was carrying out the operation or assisting a higher surgical trainee were included. All the POSSUM data were collected by D. M. at the time of surgery and stored on an EXCEL (Microsoft Corporation, Redmond, WA, USA) spreadsheet. POSSUM was not used preoperatively to assess probable risk or facilitate patient selection for surgery.

Physiological and operative factors were scored using the POSSUM method (Tables 1,2). Each factor is allocated a score of 1, 2, 4 or 8. The PS and the OS are summed separately to provide a single PS and a single OS.

The previously derived regression equation for V-POSSUM is

$$\ln(R/1-R) = (-8.0616 + (0.1552\times\text{PS}) + (0.1238\times\text{OS})),$$

where R is the mortality risk.[3]

Using the scores and regression equation for V-POSSUM, a predicted probability of death can be calculated and compared with the actual number of deaths. The primary outcome measure is 30-day mortality and this was recorded by the surgeon either at the time of death or by review of hospital records if the patient had been discharged. Individual V-POSSUM variables were also analysed for trends over time.

Over the 128-month time period, these procedures were carried out in 110 months. The data were aggregated into 12-month periods (except for the last period that had only 8 months). The ratios of observed deaths to expected deaths were used to report trends over the 11 time periods. A cumulative result of observed minus expected deaths over time chart was used to display individual results for all cases.

RESULTS

A total of 454 major vascular procedures were carried out. The detailed breakdown of procedures is shown in Table 3. Using the POSSUM classification for age, there were 82 patients aged 60 years or less (scored 1), 137 patients aged 61–70 years (scored 2) and 235 patients more than 71 years (scored 4), reflecting a typical elderly group with significant vascular disease. Using the POSSUM classification of mode of procedure, there were 339 elective procedures (scored 1), 67 urgent procedures (scored 4) and 48 emergency procedures (scored 8) (Table 1).

The number of procedures in each month ranged from 1 to 15, with a mean of 4.1 per month. There were slightly more procedures during the second to fifth periods (1996–1999) and slightly fewer during the last 3 years. There were 60 actual deaths or an overall rate of 13.2%. The number of deaths per month ranged from 0 to 3, with a mean of 0.55. There was no relation between the number of procedures carried out per month and mortality, but there was a significant decline in the annual mortality rate as shown in Table 4 ($P = 0.0002$).

The predicted number of deaths calculated from the V-POSSUM score shows that the overall predicted mortality rate was 17.2%, and that the rate varied with the 12-month period, with a high of 23.9% and a low of 9.2%. The downward trend in the predicted rate shows that the patient risk factors have changed over time and that the risk of dying has declined by almost 60% (from 23.9 to 9.2%). The factors that changed were the PS for blood pressure, pulse and cardiac status. There were also fewer multiple procedures and emergency resuscitations after the first 3 years. The overall changes in the individual scores resulted in the total PS declining from 26 to 21 ($P = 0.007$), but no significant change in the OS, from 16 to 15 (Table 5).

Table 1. The V-POSSUM physiological score

Variable	One	Two	Four	Eight
Age (years)	<60	61–70	>71	—
Cardiac status	Normal	Cardiac drugs or steroids	Oedema, warfarin, borderline cardiomegaly	Increased JVP, cardiomegaly
Respiratory status	Normal	Dyspnoea on exertion; mild COPD	Limiting dyspnoea (one flight); moderate COPD	Dyspnoea at rest (RR > 30/min)
Systolic BP (mmHg)	110–130	131–170 100–109	>170 90–99	<90
Pulse (b.p.m)	50–80	81–100 40–49	101–120	>120 <40
Glasgow coma score	15	12–14	9–11	<9
Haemoglobin (g/100 mL)	13–16	11.5–12.9 16.1–17.0	10.0–11.4 17.1–18.0	<10.0 >18.0
White cell count ($\times10^{12}$)	4–10	10.1–20.0 3.1–3.9	>20.0 <3.1	—
Urea (mmol/L)	<7.5	7.5–10.0	10.1–15.0	>15.0
Sodium (mmol/L)	>136	131–135	126–130	<126
Potassium (mmol/L)	3.5–5.0	3.2–3.4 5.1–5.3	2.9–3.1 5.4–5.9	<2.9 >5.9
Electrocardiogram	Normal	—	Atrial fibrillation (rate 60–90)	Any other abnormality

BP, blood pressure; COPD, chronic obstructive pulmonary disease; JVP, jugular venous pressure; POSSUM, Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity; RR, respiration rate; —, not applicable.

Table 2. The V-POSSUM operative severity score

Variable	One	Two	Four	Eight
Operative severity	Minor	Intermediate	Major	Major+
Multiple procedures	1	—	2	>2
Total blood loss (mL)	<100	101–500	501–999	>1000
Peritoneal soiling	None	Minor (serous fluid)	Local pus	Free bowel content, pus or blood
Presence of malignancy	None	Primary only	Nodal metastases	Distant metastases
Mode of surgery	Elective	—	Emergency resuscitation of >2 h possible; operation < 24 h after admission	Emergency (immediate surgery <2 h needed)

POSSUM, Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity; —, not applicable.

Table 3. Operative procedures

81	Elective aortic procedures
66	Emergency or acute aortic procedures
63	Carotid procedures
106	Infrainguinal arterial bypasses
83	Miscellaneous revascularization procedures, such as extra-anatomic bypass and endarterectomy
29	Other procedures
26	Upper limb procedures

Overall, the mortality rate was 23.5% lower than the predicted rate from V-POSSUM, but this improved rate can only be seen from the third period onwards, where the observed number of deaths was less than the predicted number. However, none of these differences is statistically significant (Table 4).

The improvement in the observed mortality was large (21 down to 5%), but after adjusting for the V-POSSUM scores, the improvement attributable to the surgeon was approximately 50% (the observed over expected ratios averaged approximately 50% for periods 8–11).

The results can also be used to assess whether there is a learning curve. Figure 1 shows the cumulative observed minus predicted deaths, based on the V-POSSUM predicted values. Although the patterns are not completely clear, the plot suggests that there was an excess mortality of nearly five deaths by the 80th patient and then there was a continuous improvement, or it could be suggested that for the first 200 cases, the mortality rate was as predicted and then improvement occurred. Either interpretation suggests that a possible significant change occurred after either 80 or 200 procedures.

DISCUSSION

We have reported 30-day mortality figures for an individual surgeon practising vascular surgery and showed an acceptable level of practice based on a comparison of observed with predicted death rates. The figures have also shown changes over time with a trend towards lower death rates and improved performance. Although the overall predicted 30-day mortality might seem high at 17.8%, it must be remembered that this series includes all patients and within the whole group, there were 66 emergency aortic procedures.

The V-POSSUM scores have been able to track changes in the patient casemix, which will alter expected mortality. As the surgeon's practice has changed over time, so has the observed and predicted death at each time period. The ability to track these changes enables reliable presentation of performance data on 30-day mortality. The practice of many surgeons may change over time either because the surgeon alters the casemix of patients selected for surgery or because the group served alters. V-POSSUM enables performance data to be dynamic and track these changes in a way that crude mortality rates do not.

The observed mortality rates showed a significant improvement over time, greater than 50%. Most of this was because of the patient risk factors changing (blood pressure, pulse and cardiac

Table 4. Observed and predicted mortality rates by 12-month period

12-Month period	No. patients	Deaths	Rate (%)	Predicted deaths	Predicted rate (%)	Observed rate/ predicted rate × 100
1	39	8	20.5	7	17.9	114.8
2	57	12	21.1	12	20.7	101.7
3	64	14	21.9	15	23.9	91.6
4	56	6	10.7	9	16.8	63.8
5	58	6	10.3	10	17.4	59.4
6	37	2	5.4	4	11.4	47.3
7	30	5	16.7	6	19.1	87.1
8	32	2	6.3	6	17.3	36.1
9	24	3	12.5	3	12.0	104.5
10	42	2	4.8	5	11.5	41.3
11	15	0	0.0	1	9.2	0.0
Total	454	60	13.2	78	17.2	76.8

Table 5. Summary of PS and OS by 12-month period

12-Month period	No. patients	Mean PS	Mean OSS
1	39	25.7	15.3
2	57	25.0	16.5
3	64	26.5	17.6
4	56	25.4	14.6
5	58	24.3	15.6
6	37	23.7	13.8
7	30	24.3	17.7
8	32	25.1	15.2
9	24	23.8	15.3
10	42	22.1	15.5
11	15	21.0	14.1
Total	454	24.6	15.7

OS, operative score; OSS, operative severity score; PS, physiological score.

signs) and two minor changes in operative factors, although these did not change the OS. Over the 11-year period, the observed over expected ratio was 77%, or 23% less deaths than expected. However, the performance changed over time, with performance being 50% better at the end of the study. The drop in both observed and predicted deaths is probably multifactorial. A gradual reduction in the number of emergency vascular procedures as the number of consultant vascular surgeons increased in the UK unit, followed by a switch to a smaller population in New Zealand, would lower the number of higher risk emergency procedures. It is probable that the surgeons' selection of patients has matured, such that patients with improved risks of surviving are being selected, thus lowering both observed and predicted deaths. Surgical technique and postoperative care may also have improved, possibly reflecting a learning curve since specialist appointment.

Observed over predicted ratios can be used to make comparisons with other units. In this case, the V-POSSUM parameters can estimate the mortality rates if the original surgeons, who contributed to the training set of data from which the V-POSSUM regression equation was derived, were to carry out these procedures. For the first three periods, the observed and predicted deaths were similar, but then the observed deaths were consistently lower than the predicted. Furthermore, the difference was because of the V-POSSUM model predicting high mortality rates for those with the greatest risk. It is not easy to identify whether this signifies better performance for high-risk patients by D. M. or whether the model is incorrect. Operating on high-risk patients leaves a surgeon open to criticism, especially if only crude mortality rates are reported when results will be inferior when compared with low-risk patients. V-POSSUM by adjusting for the poorer health of these patients provides protection for the surgeon in these circumstances.

Evaluation of performance may become an increasingly important issue in the future, with shorter training times, because of hours of work restrictions and narrower breadth of training, with appointment to a consultant post at an earlier stage. Although V-POSSUM has been developed specifically for vascular surgery, there are other regression equations, such as POSSUM and P-POSSUM (Portsmouth predictor equation for mortality), which apply across the range of general surgery, but require different methods of analysis.[4–8] One concern with the POSSUM dataset is that it requires human input and a degree of interpretation in some areas. There are also 18 different parameters to score, which can take 3–5 min and may inhibit its use. A recently reported minimum dataset for vascular surgery (Vascular Biochemistry and Haematology Outcome Models) only uses information obtainable before operation and would be easier to collect routinely than V-POSSUM.[9–11] This has several advantages if supported by further studies. It would enable collection of data in 1–2 min and this could be automated from computer records, which would facilitate regular updating of performance. It can also be used for non-operative patients, a group not included in most analyses. The principles in this paper, with some adjustments, are generalizable to the whole of general surgery.

Although mortality is important, it is only one aspect of surgical performance. Morbidity, patient satisfaction and assessment of the surgical team through procedure-specific outcomes are equally relevant. Mortality outcomes are not necessarily a reflection of just the operating surgeon's skill, but also rely on input from other specialties, such as intensive care, cardiology and anaesthetics.[12]

Individual evaluation of surgical performance is straightforward and has many potential benefits for surgeons, patients and funding bodies. Although often cool to the concept, surgeons should embrace the development of evaluation tools, such as V-POSSUM, and use them for the mutual benefit of themselves and their patients.

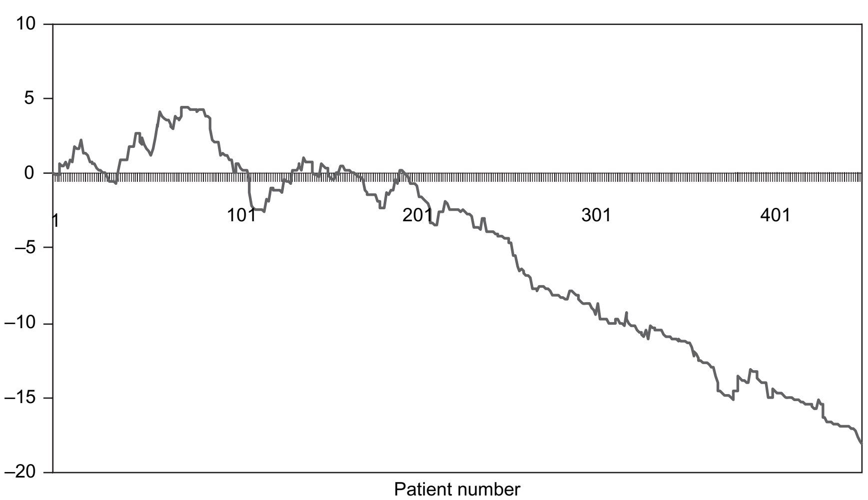

Fig. 1. Cumulative sum plot of observed minus expected deaths, 1995–2006.

REFERENCES

1. Royal Australasian College of Surgeons (RACS). *CPD Programme Information Manual 2004-2006. A Guide by the Royal Australasian College of Surgeons*, 7th edn. Melbourne: RACS, 2003.
2. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br. J. Surg*. 1991; **78**: 355–60.
3. Neary WD, Heather BP, Earnshaw JJ. The Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM). *Br. J. Surg*. 2003; **90**: 157–65.
4. 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.
5. Wijesinghe LD, Mahmood T, Scott DJA *et al*. Comparison of POSSUM and the Portsmouth predictor equation for predicting death following vascular surgery. *Br. J. Surg*. 1998; **85**: 209–12.
6. 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.
7. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. POSSUM and Portsmouth POSSUM for predicting mortality. *Br. J. Surg*. 1998; **85**: 1217–20.
8. Midwinter MJ, Tytherleigh M, Ashley S. Estimation of mortality and morbidity risk in vascular surgery using POSSUM and the Portsmouth predictor equation. *Br. J. Surg*. 1999; **86**: 471–4.
9. Tang T, Walsh SR, Prytherch DR *et al*. VBHOM, a data economic model for predicting the outcome after open abdominal aortic aneurysm surgery. *Br. J. Surg*. 2007; **94**: 717–21.
10. Prytherch DR, Ridler BMF, Ashley S; Audit and Research Committee of the Vascular Society of Great Britain and Ireland. Risk-adjusted predictive models of mortality after index arterial operations using a minimal data set. *Br. J. Surg*. 2005; **92**: 714–18.
11. Prytherch DR, Sirl JS, Weaver PC, Schmidt P, Higgins B, Sutton GL. Towards a national clinical minimum data set for general surgery. *Br. J. Surg*. 2003; **90**: 1300–305.
12. Wright J, Bradley C, Sheldon T, Lilford R. Trial by media: dangers of misinterpretation of medical statistics. *Lancet* 2006; **367**: 1139–40.