

# Perioperative mortality and morbidity prediction using POSSUM, P-POSSUM and APACHE II in Chinese gastric cancer patients: surgical method is a key independent factor affecting prognosis

Yantian Fang · Chunhsien Wu · Xiaodong Gu ·  
Zhengyang Li · Jianbin Xiang · Zongyou Chen

Received: 28 October 2012 / Accepted: 12 January 2013 / Published online: 9 March 2013  
© Japan Society of Clinical Oncology 2013

## Abstract

**Introduction** Gastric cancer is the fourth most common cancer worldwide. Predicting morbidity and mortality is important in deciding timing of surgery and type of surgery offered. APACHE II, POSSUM, and P-POSSUM are the most reliable scoring methods in use today. This is the first paper to evaluate the utility of all three scoring systems in China.

**Methods** We collected data on 851 patients (583 male and 268 female) who underwent surgery between 1991 and 2011. Physiological and pathological data was entered in spreadsheet format and analyzed using STATA version 11.0 to generate ROC curves for each scoring system.

**Results** In predicting mortality, P-POSSUM and POSSUM were most effective and APACHE II was ineffective. POSSUM predicted a higher morbidity risk than was actually encountered. Age and type of operation were found to be independent risk factors for mortality.

**Discussion** The utility of the APACHE II score in gastric cancer patients is limited. APACHE II is suitable for considering group versus individual effect. The POSSUM score is useful in general surgery, but needs improvement. We found the P-POSSUM score to be superior for morbidity and mortality prediction. P-POSSUM is useful for both the

general population and for a specific cohort. The type of surgery is a key decision point for surgeons, and independently affects prognosis. Based upon these findings and clinical scoring systems, clinicians can develop individualized treatment algorithms.

**Keywords** Gastric cancer · POSSUM · P-POSSUM · APACHE II

## Introduction

Gastric cancer is the fourth most common cancer worldwide. In Japan, the newest reported age-standardized mortality rates are 20.5 % in males and 8.0 % in females by the International Agency for Research on Cancer. In Korea, these numbers are relatively lower: 13.8 % in males and 7.3 % in females [1]. In China, the mortality rate due to gastric cancer is 40.8/100,000 and 18.6/100,000 in males and females, respectively.

The decision to perform surgery or administer chemoradiotherapy is based upon cancer stage at presentation; comprehensive treatment based on surgery is considered the standard. Outcomes depend not only on surgical expertise, but also on patient physiology, tumor pathology, and nutritional status. Therefore, early prognosis is essential. Predicting morbidity and mortality is critical in decision making with regards to timing and type of surgery.

Several scoring systems currently exist for prediction of morbidity and mortality throughout surgical disciplines. Of these, APACHE II, POSSUM, and P-POSSUM have been shown to be the most reliable and useful. APACHE II has been validated in several clinical trials [2], and estimates ICU mortality based upon laboratory values and patient characteristics, including acute and chronic disease.

Y. Fang and C. Wu contributed equally.

Y. Fang · C. Wu · X. Gu · Z. Li · J. Xiang (✉) · Z. Chen (✉)  
Department of General Surgery, Huashan Hospital  
affiliated to Fudan University, No. 12 Wulumuqizhong Rd,  
Shanghai 200040, People's Republic of China  
e-mail: xjbzhw@163.com

Z. Chen  
e-mail: czysghs@126.com

POSSUM, which was introduced in 1991 [3], was found to overestimate mortality risk; in 1996, P-POSSUM was developed, an equation which used similar physiological and operative factors but incorporated linear analysis [4].

There is an abundance of studies that address surgical morbidity and mortality. However, these data are inconclusive for patients of Asia–Pacific descent; racial and regional differences may exist. For example, when comparing English and American patients, the mortality rate predicted by POSSUM was higher than the actual mortality rate in American patients [5]. To date, few authors have evaluated gastric cancer in China using all three scoring systems. We analyzed 851 patients who presented with gastric cancer between 1991 and 2011, and found P-POSSUM to be most accurate. In addition, type of surgery independently predicts morbidity and mortality.

## Patients and methods

### Data collection and processing

Through an independent review by two groups of researchers specializing in gastric cancer, all clinical data was collected from 851 patients (583 male, 268 female, ratio 2.18:1) who presented between 1991 and 2011 at Huashan Hospital, China. Patients who were not deemed to be surgical candidates, or those who declined surgery, were excluded from the analysis.

All patients underwent gastric resection and pathological staging according to current NCCN guidelines by our institutional Department of Pathology. Patient demographic data are presented in Table 1, and were obtained from patient medical records and nursing notes. Operative data were obtained from operative, pathological, and anesthetic records, with 30-day morbidity and mortality considered to be standard entry information. All data was entered in a spreadsheet format using Microsoft Excel 2007. Statistical analysis was performed using STATA version 11.0.

### Scoring systems

APACHE II consists of three parts: acute physiology score (APS), age, and chronic physiology score (CPS). APS includes 12 parameters, ranging from 0 to 4 points, with a total score of 0–60 points. Age scores range from 0 to 6 points. CPS scores range from 2 to 5 points. In total, APACHE ranges from 0 to 71 points.

APACHE II uses the following formula to calculate the risk of death ( $R1$ ):  $\ln(R1/1 - R1) = -3.517 + (\text{APACHE II score} \times 0.146) + (0.603, \text{ if an emergency operation}) + (\text{CPS})$ .

**Table 1** Data classification method, and the number of each classification

Patient characteristics	Number	Percentage	Actual morbidity	Actual mortality
<b>Gender</b>				
Male	583	68.51	28.13	3.77
Female	268	31.49	30.22	2.24
<b>Age</b>				
≤50 years	186	21.86	17.20	1.61
50–70 years	445	52.29	26.07	1.12
≥71 years	220	25.85	44.09	9.09
<b>PTNM classification</b>				
T1	96	11.28	28.13	2.08
T2	122	14.34	29.51	0.82
T3	456	53.58	31.36	2.63
T4	177	20.80	22.03	7.34
N0	287	33.73	27.18	1.39
N1	282	33.14	31.21	2.84
N2	109	12.81	31.19	0.92
N3	173	20.33	26.01	8.67
M0	622	73.09	28.62	2.25
M1	229	26.91	29.26	6.11
<b>Pathological stage</b>				
I	184	21.62	27.17	1.09
II	313	36.78	29.07	2.56
III	125	14.69	28.80	2.40
IV	229	26.91	29.69	6.55
<b>Pyloric obstruction</b>				
No	735	86.37	29.12	0.64
Yes	116	13.63	26.72	4.47
<b>Surgical methods</b>				
Proximal gastrectomy	34	4.00	35.29	0.00
Distal gastrectomy	466	54.76	29.18	1.50
Total gastrectomy	137	16.10	28.47	2.92
Combined organ resection	103	12.10	33.01	4.85

Actual mortality and complication rates are also listed

The POSSUM score has two main components, physiological and operative severity, which are calculated to predict postoperative prognosis. We used POSSUM as originally defined by Copeland et al. [3], and P-POSSUM as defined by Whiteley et al. [4].

The following equations were used to calculate morbidity and mortality for POSSUM ( $R2 = \text{mortality}$ ,  $R3 = \text{morbidity}$ ).

$$\text{Log}_e \frac{R2}{(1 - R2)} = -7.04 + (0.13 \times \text{physiological\_score}) + (0.16 \times \text{operative\_severity\_score})$$

$$\text{Log}_e \frac{R3}{(1 - R3)} = -5.91 + (0.16 \times \text{physiological\_score}) + (0.19 \times \text{operative\_severity\_score})$$

We used the following equation to calculate mortality for P-POSSUM ( $R4$  = risk of mortality).

$$\text{Log}_e \frac{R4}{(1 - R4)} = -9.065 + (0.1692 \times \text{physiological\_score}) + (0.1550 \times \text{operative\_severity\_score}).$$

### Evaluation of scoring systems

We calculated each patient's score according to each of the three described systems. We generated a mortality observed/expected (O/E) ratio and morbidity O/E ratio for each scoring system. Receiver operating characteristic (ROC) curves were generated to demonstrate the sensitivity and specificity of morbidity as predicted by POSSUM. The area under the curve determines accuracy. Regression analysis was used to evaluate the effectiveness of each scoring system.

### Statistical methods

Student's *t* test was used to compare categorical and continuous variables, and to analyze both observed and predicted outcome measures for each scoring system. A *p* value of less than 0.05 was considered statistically significant. Statistical analysis was performed using STATA software, version 11.0.

## Results

### Mortality rates as predicted by APACHE II, POSSUM, and P-POSSUM scoring systems

Both APACHE II and POSSUM estimations of preoperative mortality risk were generally high. However, POSSUM had an acceptable specificity (false-negative) rate. P-POSSUM overestimated mortality risk, but was the most accurate of the three systems, with a false-negative rate of 0 (Table 2).

### ROC analysis

We calculated the scores for each patient with these three forecasting methods from the patients' information we collected. We compared the value from the prediction

**Table 2** O/E ratios of three scoring systems for postoperative mortality

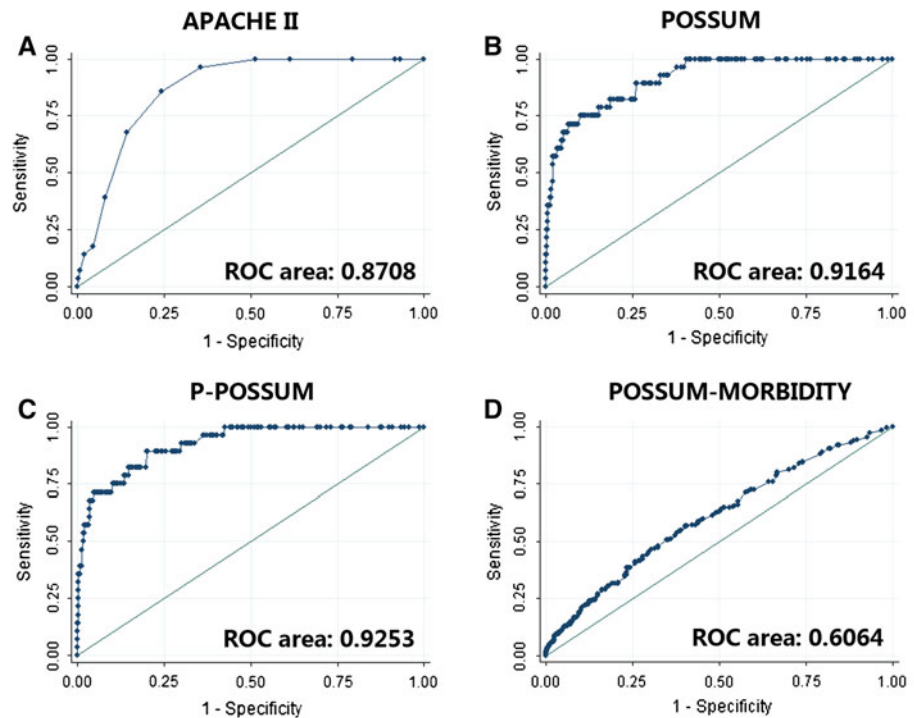
Patient characteristics	APACHE II		POSSUM		P-POSSUM	
	Predicted ratio (%)	O/E	Predicted ratio (%)	O/E	Predicted ratio (%)	O/E
Age						
≤50 years	3.17	0.51	7.15	0.23	2.14	0.75
50–70 years	4.52	0.25	8.81	0.13	2.81	0.40
≥71 years	6.62	1.37	16.84	0.54	7.39	1.23
PTNM classification						
T1	41.38	0.05	4.70	0.44	1.30	1.00
T2	47.46	0.02	7.35	0.11	2.61	0.31
T3	48.17	0.05	9.92	0.27	3.52	0.75
T4	49.88	0.15	17.42	0.42	6.92	1.06
N0	4.61	0.30	6.17	0.23	2.00	0.70
N1	4.77	0.59	10.21	0.28	3.56	0.80
N2	4.68	0.20	10.76	0.09	3.79	0.24
N3	5.06	1.71	18.12	0.48	7.41	1.17
M0	4.65	0.48	7.44	0.30	2.48	0.91
M1	5.08	1.20	18.90	0.32	7.55	0.81
Pathological stage						
I	4.46	0.24	5.35	0.20	1.67	0.65
II	4.72	0.54	7.85	0.33	2.61	0.98
III	4.75	0.51	9.44	0.25	3.31	0.73
IV	5.09	1.29	18.93	0.35	7.58	0.86
Pyloric obstruction						
No	4.70	0.52	9.32	0.26	3.29	0.74
Yes	5.18	1.66	18.51	0.47	7.39	1.17
Surgical methods						
Proximal gastrectomy	4.69	0.00	7.79	0.00	2.55	0.00
Distal gastrectomy	4.64	0.32	8.19	0.18	2.80	0.54
Total gastrectomy	4.49	0.65	8.97	0.33	2.85	1.02
Combined organ resection	5.34	0.91	15.90	0.31	6.36	0.76
Exploratory laparotomy	5.11	2.12	18.11	0.60	7.54	1.43

formula with the actual value. Using regression analysis, the three scoring systems were found to differ with regard to mortality prediction (Fig. 1a–c). P-POSSUM was the most effective, followed by POSSUM and APACHE II.

### POSSUM predicted morbidity rates

POSSUM was found to have a higher predicted than actual morbidity risk, although it was more accurate in patients with N3 disease or who required combined organ resection.

**Fig. 1** Predicting mortality with three scoring systems. **a** Data analyzed by APACHE II; the AUC area is 0.8708. **b** Data analyzed by POSSUM; the AUC area is 0.9164. **c** Data analyzed by P-POSSUM; the AUC area is 0.9253. **d** Predicting morbidity of 851 patients with P-POSSUM. **d** The AUC area is 0.6064. Confidence intervals are listed



SCORING SYSTEMS	AUC	Std. Err.	-Asymptotic Normal- [95% Conf. Interval]		Cutoff Point	LR+	LR-
APACHE II	0.8708	0.0214	0.82878	0.91276	0.0605	3.5449	0.1884
POSSUM	0.9164	0.0239	0.86946	0.96334	0.2092	7.6204	0.2773
P-POSSUM	0.9253	0.0225	0.88120	0.96948	0.0419	4.4535	0.1340
POSSUM-MORB	0.6064	0.0214	0.56436	0.64842	0.3989	1.4091	0.7243

However, in patients aged under 50, those with N1 and N2 disease, Stage III patients, and those who underwent total gastrectomy and exploratory laparotomy, POSSUM exaggerated morbidity risk (Fig. 1d).

**Logistic regression analysis**

**Mortality**

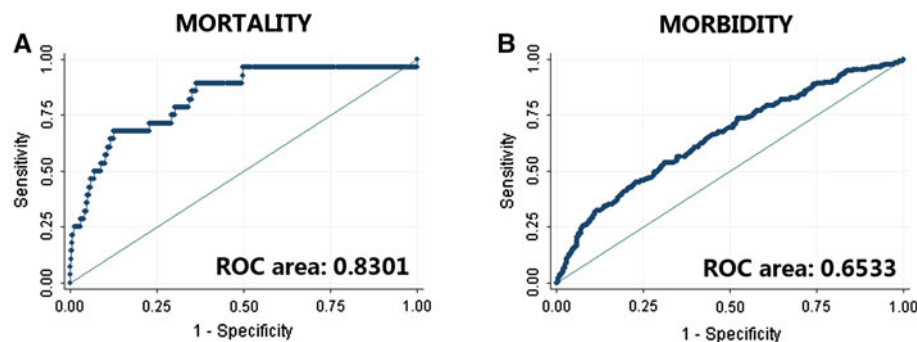
We established the following regression equation to predict mortality. We tested the goodness of fit of the equation with the method of Hosmer–Lemeshow. The result showed the *p* value was 0.1315, indicating a high goodness of fit. As the correlation between the independent variables, we chose the stepwise regression backward elimination method for variable selection. Age and surgical approach are the last two independent variables in the model. Defining whether death occurred as the dependent variable, we presented logistic regression with gender, age, and surgical methods as independent variables. As there was no case of death in the distal gastrectomy group in all the three age groups and few cases of death in the under-50 age group, we first selected the over-

50 age groups, as well as the other four surgical approach method groups for analysis. We considered the main effects of the independent variables, as there was no interaction between age and surgical approach. We confirmed that gender had no effect on whether death occurred (*p* = 0.45). If gender and surgical approach remain unchanged, the risk of death in the over-70 age group was 8.89 (95 % confidence interval 3.21, 24.63) times higher than in the 50–70-year-old group. If gender and age remain unchanged, the death risk from combined organ resection was 3.79 (95 % confidence interval 1.10, 13.04) times higher than the risk of distal gastrectomy, and the death risk from exploratory laparotomy was 6.73 (95 % confidence interval 2.31, 19.63) times higher than the risk from distal gastrectomy. There was no significant difference between the other surgical methods. We then confirmed there was no significant difference in the risk of death between the different surgery methods in the over-50 age group (Fig. 2a).

$$P' = 0.0813 \times \text{age} - 0.4239 \times T + 0.2840 \times N - 1.1935 \times M + 0.3978 \times \text{stage} + 0.4555 \times \text{po} + 0.6955 \times \text{oper} - 10.8958$$

$$R' = 1/(1 + e(-P')).$$

**Fig. 2** ROC curve of mortality and morbidity established by logistic regression analysis. **a** AUC area is 0.8301. Cutoff point is 0.0620. Positive likelihood ratio (LR+) is 5.4751. Negative likelihood ratio (LR-) is 0.3669. **b** AUC area is 0.6533. Cutoff point is 0.3155. LR+ is 1.7275. LR- is 0.6703



Logistic	AUC	Std. Err.	-Asymptotic Normal- [95% Conf. Interval]		Cutoff Point	LR+	LR-
mortality	0.8301	0.0429	0.74603	0.91418	0.0620	5.4751	0.3669
morbidity	0.6533	0.0210	0.61223	0.69440	0.3155	1.7275	0.6703

### Morbidity

We established the following regression equation to predict morbidity. We tested the goodness of fit of the equation with the method of Hosmer–Lemeshow. The result showed a *p* value of 0.1064, indicating a high goodness of fit. We chose stepwise regression backward elimination method for variable selection. Age and surgical approach are the last two independent variables in the model. If surgical approach remained unchanged, the risk of complication in the over-70 age group was 3.91 (95 % confidence interval 2.43, 6.30) times higher than in the under-50 age group, 2.22 (95 % confidence interval 1.57, 3.15) times higher than in the 50- to 70-year-old group. If the surgical approach is defined as a variable, the risk of complication in the exploratory laparotomy group was 3.12 (95 % confidence interval 1.29, 7.56) times lower than in the proximal gastrectomy group, 2.20 (95 % confidence interval 1.28, 3.80) times lower than in the distal gastrectomy group, 2.37 (95 % confidence interval 1.25, 4.49) times lower than in the total gastrectomy group, and 2.36 (95 % confidence interval 1.22, 4.58) times lower than in the combined organ resection group. There was no significant difference between the other surgical methods (Fig. 2b).

$$P' = 0.0405 \times \text{age} - 0.3206 \times T + 0.0318 \times N - 0.2392 \times M + 0.3463 \times \text{stage} + 0.2054 \times \text{po} + 0.0932 \times \text{oper} - 3.0223$$

$$R' = 1/(1 + e(-P')).$$

### Discussion

Audit is becoming increasingly more important in clinical practice and has gained attention as a tool for clinical and scientific research. APACHE II and POSSUM are audit

tools used in general surgery and several other surgical fields. In general surgery, prediction of colorectal cancer risk is the focus of research. However, few studies exist that predict risk in gastric cancer patients, especially in the Chinese and Asian populations. Various risk prediction models may provide a comparative platform that would make it possible to compare operative risk in different diseases, from different hospitals, and in different regions.

APACHE II is based upon objective physiological parameters and is less affected by treatment. Several studies report that APACHE II is closely related to disease severity; the higher the score, the more severe the disease and the higher the risk of death. An APACHE II score greater than 20 is associated with a 100 % mortality rate [6]; analysis of this research provides an overall mortality distribution.

APACHE II is quite limited, as it was designed for the ICU and is more predictive of group versus individual effect. Van Le et al. [7] confirmed that the sensitivity of the APACHE II score in the individual was only 54 %, although it was used to assess 45 gynecological patients in the ICU. Other authors considered APACHE II to be ineffective in evaluating patients under special circumstances, such as trauma, TPN, acute myocardial infarction, and severe congestive heart failure [8, 9]. In these situations, mortality was always estimated to be higher than actual when the score was low. In contrast, the predicted mortality rate was always lower than actual when the score is high. As these articles reported, APACHE II parameters were used within 24 h after ICU admission; Knaus et al. [2] believed that parameters sampled during ICU admission would be more meaningful. Further scholars think that dynamic ratings (assessing patients over several days) may be more effective in predicting mortality. APACHE II is a general rating system that does not consider pathology; however, pathology has a great influence on the general



health of the patient. APACHE II was designed by Western scientists; thus, some regional differences may exist when compared to Asian countries.

Most scoring systems reflect vital signs and vital organ function. Certain indicators, such as the Glasgow Coma Scale, hold no special significance for patients with gastric cancer; this type of surgery is generally elective and those with poor vital signs will not be subjected to surgery. Conversely, chronic disease that may affect prognosis and physiological indicators are not included in the statistics.

In many European countries, POSSUM, one of the most important common scoring methods, has become routine in preoperative evaluation of patients. It not only evaluates the risk of surgery, but also provides an objective basis for patients to choose surgical options and strengthen intervention measures in high-risk patients. Lamb used POSSUM to demonstrate that controlling the extent of surgery, and balancing the risk of surgery and extent of dissection, reduced mortality in 180 gastric cancer patients [10]. Hartley and Sagar [11] and Sah et al. [12] showed the excellent predictive value of POSSUM through a retrospective analysis of gastrointestinal surgery patients. Similar results were shown for POSSUM by Mosquera et al. (454 vascular surgery patients) [13], Tambyraja et al. (laparoscopic cholecystectomy) [14], and Bennett-Guerro et al. (liver transplantation). However, in certain situations, this system can be problematic, such as overpredicting morbidity in low-risk groups [15]. Das et al. [16] analyzed 468 gynecological tumors using POSSUM, and found that predictions of operative mortality and the actual outcomes were inconsistent. Copeland et al. [3] found that the results of POSSUM were satisfactory for general surgery, but needed improvement for use in orthopedic surgery. In this case, when the POSSUM score was less than 40 %, the predicted mortality rate was greater than 100 % of the actual outcome. However, when the POSSUM score was greater than 60 %, the number of deaths was underestimated. The false-negative rate of POSSUM was low. The actual number of deaths in each group was within the range of predicted values. P-POSSUM has the best predictive value of all three scoring methods, and is an improvement over the POSSUM scoring system. P-POSSUM has satisfactory results not only in the general population, but also among specific cohorts. Prytherch et al. [17] found, through the analysis of 10,000 general surgery patients, that the P-POSSUM scoring system's ability to accurately predict mortality was better than POSSUM. Jones et al. then did a comparative study of surgical risk in individual surgical ICU patients, using POSSUM and APACHE II, and the AUC curve for statistical analysis (AUC, POSSUM = 0.75; AUC, POSSUM morbidity = 0.82; AUC, APACHE II = 0.54) [18]. Thus, the ability of POSSUM to accurately predict outcome is

significantly higher than APACHE II. De Cassia Braga Ribiero and Kowalski analyzed postoperative morbidity and mortality after pharyngeal cancer surgery, using APACHE II, POSSUM, and the ASA scoring systems [19]. The predictive value was graded as follows: POSSUM > APACHE II > ASA. In this case, POSSUM and P-POSSUM were better able to predict mortality, but POSSUM was weak in predicting complications.

In statistical analysis, POSSUM has several drawbacks. Electrocardiogram grades do not accurately reflect patient risk. Because of the increased likelihood of ventricular arrhythmias, myocardial ischemia impacts both morbidity and mortality rates, but these are not reflected in the ECG grading standards in the POSSUM scoring system. Subjective elements easily cause bias. Intraoperative blood loss and abdominal contamination cannot be precisely determined. The surgical acuity score therefore has too many subjective factors. Some patients have several physiological indicators corrected before surgery, such as anemia, malnutrition, and electrolyte balance disorders. Some authors therefore suggest that risk assessment for gastrointestinal surgery should be conducted at two points: 24 h after admission, and postoperative day one. The first assessment is to determine whether the patient will tolerate surgery and to determine the timing and type of operation. The second assessment is to predict morbidity and mortality and to guide further treatment.

With logistic regression analysis, we found only age and type of surgery to be independent factors of patient death. Similarly, age, characteristics of the primary tumor, and type of surgery are independent predictors of complications. Although TNM is an independent factor affecting prognosis in gastric cancer, survival decreases with increased TNM stage; we did not find TNM staging to be an independent predictor of postoperative mortality.

The selection of type of operation that may carry an increased risk of morbidity and mortality is a critical issue for surgeons. Westerners often have associated obesity, circulatory, and pulmonary comorbidities; thus, European and American surgeons seldom perform D2 radical gastric cancer resection. In contrast, D2 radical gastric cancer resection has long been considered a standard surgical procedure in Japan. Combined gastrectomy often includes the spleen, pancreas, colon, and duodenum. Although some of these radical procedures do not demonstrate evidence of increased long-term survival, a more aggressive operation may significantly impact postoperative morbidity and mortality rates. Therefore, surgeons should aim for a radical cure with as much of a controllable risk as possible. On the other hand, as combined organ resection does not carry a significant mortality risk in patients with advanced disease, if resectable, surgeons should aim for as complete an operation as possible, reducing tumor burden.

## Conclusion

As gastrointestinal surgery advances, selecting a personalized treatment based on patient condition is becoming increasingly important. Preoperative assessment for Chinese patients needs improvement. The impact of gastric cancer on patients with perioperative risk needs to be balanced with current scoring systems so that an appropriate formula can be used for Chinese patients. Clinicians may develop individualized treatment plans according to these formulas, selecting the appropriate surgical approach, enhancing preoperative organ protection, and optimizing postoperative treatment.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. Ferlay J, Shin HR, Bray F et al (2010) GLOBOCAN 2008 v2.0, cancer incidence and mortality worldwide: IARC CancerBase No. 10 [Internet]. International Agency for Research on Cancer, Lyon. <http://globocan.iarc.fr> Accessed 4 Dec 2012
2. Knaus WA, Draper EA, Wagner DP et al (1985) APACHE II: a severity of disease classification system. *Crit Care Med* 13: 818–829
3. Copeland GP, Jones D, Walters M (1991) POSSUM: a scoring system for surgical audit. *Br J Surg* 78:355–360
4. Whiteley MS, Prytherch DR, Higgins B et al (1996) An evaluation of the POSSUM surgical scoring system. *Br J Surg* 83:812–815
5. Bennett-Guerrero E, Hyam JA, Shaefi S et al (2003) Comparison of P-POSSUM risk-adjusted mortality rates after surgery between patients in the USA and the UK. *Br J Surg* 90:1593–1598
6. Bosscha K, Reijnders K, Hulstaert PF et al (1997) Prognostic scoring systems to predict outcome in peritonitis and intra-abdominal sepsis. *Br J Surg* 84:1532–1534
7. Van Le L, Fakhry S, Walton LA et al (1995) Use of the APACHE II scoring system to determine mortality of gynecologic oncology patients in the intensive care unit. *Obstet Gynecol* 85:53–56
8. Ridley S (1998) Severity of illness scoring systems and performance appraisal. *Anaesthesia* 53:1185–1194
9. Ip SP, Leung YF, Ip CY et al (1999) Outcomes of critically ill elderly patients: is high-dependency care for geriatric patients worthwhile? *Crit Care Med* 27:2351–2357
10. Lamb P, Sivashanmugam T, White M et al (2008) Gastric cancer surgery—a balance of risk and radicality. *Ann R Coll Surg Engl* 90:235–242
11. Hartley MN, Sagar PM (1994) The surgeon's 'gut feeling' as a predictor of post-operative outcome. *Ann R Coll Surg Engl* 76:277–278
12. Sah BK, Zhu ZG, Chen MM et al (2009) Effect of surgical work volume on postoperative complication: superiority of specialized center in gastric cancer treatment. *Langenbecks Arch Surg* 394:41–47
13. Mosquera D, Chiang N, Gibberd R (2008) Evaluation of surgical performance using V-POSSUM risk-adjusted mortality rates. *ANZ J Surg* 78:535–539
14. Tambyraja AL, Kumar S, Nixon SJ (2005) POSSUM scoring for laparoscopic cholecystectomy in the elderly. *ANZ J Surg* 75:550–552
15. Bennett-Guerrero E, Feierman DE, Barclay GR et al (2001) Preoperative and intraoperative predictors of postoperative morbidity, poor graft function, and early rejection in 190 patients undergoing liver transplantation. *Arch Surg* 136:1177–1183
16. Das N, Talaat AS, Naik R et al (2006) Risk adjusted surgical audit in gynaecological oncology: P-POSSUM does not predict outcome. *Eur J Surg Oncol* 32:1135–1138
17. Prytherch DR, Sutton GL, Boyle JR (2001) Portsmouth POSSUM models for abdominal aortic aneurysm surgery. *Br J Surg* 88:958–963
18. Jones DR, Copeland GP, de Cossart L (1992) Comparison of POSSUM with APACHE II for prediction of outcome from a surgical high-dependency unit. *Br J Surg* 79:1293–1296
19. de Cassia Braga Ribeiro K, Kowalski LP (2003) APACHE II, POSSUM, and ASA scores and the risk of perioperative complications in patients with oral or oropharyngeal cancer. *Arch Otolaryngol Head Neck Surg* 129:739–745