University of Groningen

Comparison of different risk-adjustment models in assessing short-term surgical outcome after transthoracic esophagectomy in patients with esophageal cancer

Bosch, D.J.; Pultrum, B.B.; de Bock, Gertruida H; Oosterhuis, J.K.; Rodgers, M.G.G.; Plukker, John T.H.M.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2011

Link to publication in University of Groningen/UMCG research database

Clinical Science

Comparison of different risk-adjustment models in assessing short-term surgical outcome after transthoracic esophagectomy in patients with esophageal cancer

Dirk J. Bosch, B.Sc.[a], Bastiaan B. Pultrum, M.D., Ph.D.[a],
Gertrude H. de Bock, Ph.D.[b], Jurjen K. Oosterhuis, M.D.[c],
Michael G. G. Rodgers, M.D.[d], John T. M. Plukker, M.D., Ph.D.[a,*]

[a]Department of Surgery/Surgical Oncology, [b]Department of Epidemiology, [c]Department of Anesthesiology, and [d]Intensive Care Unit, University Medical Center Groningen, University of Groningen, Hanzeplein 1, PO Box 30001, 9700 RB Groningen, The Netherlands

Abstract

BACKGROUND: Different risk-prediction models have been developed, but none is generally accepted in selecting patients for esophagectomy. This study evaluated 5 most frequently used risk-prediction models, including the American Society of Anesthesiologists, Portsmouth-modified Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (P-POSSUM), and the adjusted version for Oesophagogastric surgery (O-POSSUM), Charlson and the Age adjusted Charlson score to assess postoperative mortality after transthoracic esophagectomy.

METHODS: Data were obtained from 278 consecutive esophageal cancer patients between 1991 and 2007. Performance in predicting postoperative mortality (in-hospital and 90-day mortality) were analyzed regarding calibration (Hosmer and Lemeshow goodness-of-fit test) and discrimination (area under the receiver operator curve).

RESULTS: The Hosmer and Lemeshow goodness-of-fit test was applied to each model and showed a significant outcome for only the P-POSSUM score ($P = .035$). The receiver operator curve indicated discriminatory power for P-POSSUM (.766) and for O-POSSUM (.756), other models did not exceed the minimal surface of .7.

CONCLUSIONS: Postoperative mortality after esophagectomy was best predicted by O-POSSUM. However, it still overpredicted postoperative mortality.

© 2011 Elsevier Inc. All rights reserved.

Esophageal cancer is associated with high rates of perioperative morbidity and mortality and a relatively low overall 5-year survival rate of approximately 25%.[1] The incidence is increasing rapidly and appears to be most prominent in vulnerable and fragile elderly older than age 70 years who withstand major surgical insult as well.[2] Unfortunately, many elderly patients have serious comorbidities interfering with the outcome of treatment.[3,4] Careful preoperative assessment of fitness and subsequent selection of appropriate surgical candidates are important steps improving short-term outcome for individuals who undergo an esophagectomy.

New standard treatment methods, including neoadjuvant chemoradiation with reported complete responses of 20% to

* Corresponding author. Tel.: +31-50-361 23 17; fax: +31-50-361 48 73.
E-mail address: j.th.plukker@chir.umcg.nl

40% after resection, can be performed safely in a great part of these patients.[5] Nevertheless, surgery remains the primary curative option.[6,7] However, esophagectomy as a high-risk complex surgical procedure has a severe postoperative complication rate of up to 50% with a relatively high postoperative mortality rate of around 5%, and in some cohorts approaching 10% to 15%.[6,8,9]

Preoperative risk stratification for postoperative mortality may help patients and families address the magnitude of both the disease and the therapy. It is pivotal for both the patient and the surgeon to realistically assess the magnitude of the surgical insult. Therefore, we propose to assess several preoperative scoring systems that have each been validated as predictive of severe postoperative morbidity and mortality.[10–14]

These risk stratification/adjustment systems include the Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM), its Portsmouth (P-POSSUM) and O-POSSUM modifications, the Charlson Comorbidity Index (CCI) with the Age-Adjusted Charlson Score (ACCI) version and the standard American Society of Anesthesiologists (ASA) classification systems. In most of these systems age was not included as a dominant predictor of morbidity that is uniquely relevant to esophageal cancer presenting nowadays in more aging patients.[14] Until now, there were no published studies comparing all 5 of these comorbidity models (P-POSSUM, O-POSSUM, Charlson, ACCI, and ASA score) for patients after esophagectomy. We examined which of these 5 most frequently used comorbidity models could predict short-term surgical outcomes accurately after curatively intended resection in esophageal cancer patients.

Patients and Methods

Patient characteristics

Between January 1991 and December 2007 there were 280 consecutive patients with cancer of the esophagus who underwent a surgical resection with curative intent. Two patients with missing follow-up data were excluded from the analysis. In the remaining group of 278 patients an analysis was performed based on prospectively registered data from a computerized database of all esophageal procedures at our university hospital (Table 1). Data from this study were evaluated according to the rules of the ethical board at our institute. There were no systemic changes over the study period in the methods of acquiring patient comorbidity data.

Preoperative work-up

Preoperative evaluation consisted of physical examination, standard laboratory tests, and detailed preoperative risk assessments. Staging was performed by endoscopic ultra-

Table 1 Variables of P-POSSUM and O-POSSUM

Physiological score	Surgical severity score
Age, y	Surgical severity
Glasgow Coma Score	Multiple procedures
Cardiac signs	Total blood loss*
Respiratory signs	Peritoneal soiling*
Electrocardiography	Presence of malignancy*
Systolic pressure, mm Hg	Mode of surgery
Pulse rate, beats/min	
Hemoglobin level, g/dL	
White blood cell count, $\times10^{12}$/L	
Urea level, mmol/L	
Sodium level, mmol/L	
Potassium level, mmol/L	

*Only in P-POSSUM.

sonography with fine-needle aspiration of suspected lesions and 16- to 64-slice multidetector computed tomography of the chest, abdomen, and cervical region. From 1996 on, all patients diagnosed as T3-4 or N1 were additionally staged with ^{18}F-fluoro-2-deoxy-D-glucose positron emission tomography (PET) and PET/computed tomography fusion was applied in case of anatomic difficulties on PET assessment. Since 2007, neoadjuvant chemoradiation consisting of paclitaxel 50 mg/m^2 and carboplatin (area under the curve [AUC], 2) on days 1, 8, 15, 22, and 29, with concurrent radiotherapy of 41.4 Gy (23 fractions of 1.8 Gy), was administered to 10 patients as part of a randomized control trial with surgery alone.

Surgery

Surgery in our tertiary referral center was performed by 2 experienced surgeons. All patients underwent a curative intended open radical transthoracic esophagectomy consisting of a subtotal esophageal resection including a 2-field lymphadenectomy of nodes at the celiac trunk, along the upper border of the pancreas, para-aortic region, and mediastinal nodes. Pathologic staging was based according to the latest edition of the TNM classification for esophageal cancer.[15]

Comorbidity and mortality indexes

Overall comorbidity severity was classified according to the modified P-POSSUM, O-POSSUM, CCI, ACCI, and ASA scores.

The original POSSUM score overpredicted mortality in low-risk patients[16] and therefore transformed into the Portsmouth predictor equation (P-POSSUM), with a different logistic regression. Both risk prediction models are based on a preoperatively available 12-factor physiological score and a 6-factor surgical severity score obtained after surgery. To fulfill the need for a specialized risk prediction model for

esophagogastric surgery, the adapted O-POSSUM equation was designed[17] (Table 1).

The 19 conditions of the CCI were found to influence survival significantly and were given a weighted, risk-adjusted comorbidity index value, varying from 1 to 6 points, for the individual patient.[18] Patients with a low score were considered to have minimal comorbid diseases in their medical history. In our study we used the modification by Romano et al[19] because it excludes cancer diagnosis in determining comorbidity and is used commonly in cancer outcomes research. The CCI reflects both the number and gravity of comorbid diseases. Besides the Charlson score, we also used the ACCI scoring system, which characterized the impact of age and comorbidity on disease progression and survival after surgery.[20] Both models initially were developed for administrative databases and not for individual patient level data sets. The commonly used ASA Physical Status classification is readily available and widely accepted to stratify surgical patients according to their perioperative risk. It varies from ASA 1 (normal healthy patient in good condition) to ASA 5 (moribund patient, not expected to survive).[21] ASA class is assigned by the attending anesthesiologist after completing a structured review of physical status just before the patient's surgical procedure. Although the ASA classification was not initially intended to predict survival beyond the perioperative period, several investigators have shown a prognostic value for the ASA classification beyond this period.[21]

Statistical analysis

The primary outcome was postoperative mortality, which we defined as death within 90 days after esophagectomy or any death during admission in the hospital where the resection was performed. This time period was applied to include all surgery-related deceased patients. The observed number of deceased patients was divided by the number of expected deceased patients and gave a standard mortality ratio (SMR). The performance of P-POSSUM and O-POSSUM, CCI, ACCI, and the ASA score in predicting postoperative mortality was analyzed regarding calibration and discrimination. Calibration refers to the agreement between observed outcomes and predicted probabilities and concerns the expected mortality rate for a group of patients. Comparison between observed and expected deaths for each model was analyzed with the Hosmer and Lemeshow (HL) goodness-of-fit test.[22,23] Higher values of the HL statistic represent poorer model calibration. In this analysis a P value of less than .05 was considered to show a statistically significant lack of fit. Discrimination refers to the ability to distinguish patients who will die from those who will survive by computing the area under the receiver operating characteristic (ROC) curve. Values between .7 and .8 suggest reasonable or moderate discrimination and values exceeding .8 suggest good or excellent discrimination.

For a better applicability in clinical practice, both POSSUM models were divided into 3 risk categories: group I (low risk), with a postoperative mortality rate of 0% to less than 8%; group II (intermediate risk), with a postoperative mortality rate of 8% to less than 15%; and group III (high risk), with a postoperative mortality rate of 15% to 100%.[24]

To counteract the possibility of changes in hidden care over the study period (1991–2007), we divided this period in 3 segments. The predictive powers of these models were analyzed in each time segment and were compared with the overall predictive power. All statistical analyses were conducted using the statistical software SPSS 16.0.2 (SPSS, Inc., Chicago, IL).

Results

The clinicopathologic characteristics of the patients are summarized in Table 2. The 90-day postoperative mortality rate was 6.5% (18 patients), including an in-hospital mortality rate of 5.4% (n = 15). The overall comorbidity severity evaluated according to the 5 most commonly used models was as follows.

Evaluation of the POSSUM equation

The expected mortality ratio by P-POSSUM was 6.2% (17 patients), with a SMR of 1.05 (18 of 17). O-POSSUM expected a postoperative mortality rate of 9.7% (27 patients), which leads to a SMR of .67 (18 of 27). This value indicates an overestimation by O-POSSUM. The risk classification of both POSSUM models, with subdivision in observed and expected mortality rates, are summarized in Table 3.

Calibration of the HL statistic showed no fit to the observed data for P-POSSUM (χ^2 = 16.580; 8 d.f. (degrees

Table 2   Patient (n = 278) and tumor characteristics according to postoperative outcome

	Postoperative survivors (n = 260), (%)	Postoperative deceased patients (n = 18), (%)
Median age, y (range)	63 (29y–85y)	70 (55y–81y)
Sex, male/female	214/46 (82.3/17.7)	15/3 (83.3/16.7)
Histology		
Adenocarcinoma	218 (83.9)	17 (94.4)
Squamous cell carcinoma	42 (16.2)	1 (5.6)
Localization		
Midesophageal	22 (8.5)	1 (5.5)
Distal esophagus	238 (91.5)	16 (94.5)
Tumor stage		
I	38 (14.7)	1 (5.6)
IIa	68 (26.2)	2 (11.1)
IIb	34 (13.1)	0 (.0)
III	107 (41.2)	14 (77.8)
IVa	13 (5.0)	1 (5.6)

Table 3   Outcomes of P-POSSUM and O-POSSUM stratified for risk groups: observed and expected mortality rates

Score (%)*	Patients, n	Observed mortality, % (n)	Expected mortality, % (n)
P-POSSUM			
0 to <8	219	3.7 (8)	3.5 (8)
8 to <15	41	12.2 (5)	10.6 (4)
≥15–100	18	27.8 (5)	29.1 (5)
Total	278	6.5 (18)	6.2 (17)
O-POSSUM			
0 to <8	137	1.5 (2)	4.8 (7)
8 to <15	97	9.3 (9)	11.4 (11)
≥15–100	44	15.9 (7)	20.8 (9)
Total	278	6.5 (18)	9.7 (27)

*Possum risk group.

Table 5   Outcomes of CCI, ACCI, and ASA score

Score	Patients, n	Observed mortality, % (n)
Charlson		
0	143	5.6 (8)
1	82	4.9 (4)
2	35	11.4 (4)
3	16	12.5 (2)
4	2	.0 (0)
ACCI		
0	20	.0 (0)
1	55	.0 (0)
2	64	7.8 (5)
3	71	7.0 (5)
4	39	12.8 (5)
5	15	.0 (0)
6	12	25.0 (3)
7	2	.0 (0)
ASA		
1	36	.0 (0)
2	177	6.2 (11)
3	60	8.3 (5)
4	5	40.0 (2)

of freedom); $P = .035$), in contrast to the calibration of O-POSSUM ($\chi^2 = 7.074$, 8 d.f.; $P = .529$; Table 4). The area under the ROC curve for P-POSSUM was .766 (95% confidence interval [CI], .67–.86; $P = .000$), indicating discriminatory power for postoperative mortality. A similar result was found for O-POSSUM, the ROC curve analysis revealed discriminatory capability for postoperative deaths with an AUC of .756 (95% CI, .67–.84; $P = .000$).

Evaluation of the CCI and ACCI scores

In our cohort the CCI score ranged from 0 to a maximum of 4 points. Patients with a CCI score of 0 points had an observed postoperative mortality of 5.6% (8 patients), patients with a score of 1 point had an observed postoperative mortality of 4.9% (4 patients), patients with a score of 2 points had an observed postoperative mortality of 11.4% (4 patients), and patients with a score of 3 points had an observed postoperative mortality of 12.5% (2 patients); there were no patients with a score of 4 points (Table 5). The ACCI score in the study group ranged from 0 to 7 points and showed similar postoperative mortality rates, with an increased risk of mortality in general with higher scores (Table 5).

The HL goodness-of-fit test, when applied to the CCI score, indicated a good fit to the observed postoperative deaths ($\chi^2 = .833$; 2 d.f.; $P = .659$), as well as the ACCI

score, which showed a similar fit to the observed data ($\chi^2 = 5.174$; 4 d.f.; $P = .270$; Table 4). The area under the ROC curve for the CCI score was .567 (95% CI, .42–.71; $P = .344$), indicating no discriminatory power. Similar results were found regarding the area under the ROC curve for the ACCI score, which showed the same poor discriminatory power (.684; 95% CI, .58–.79; $P = .009$). Because neither of the models showed a good fit with the observed data, they were not divided into risk categories.

Evaluation of the ASA score

There were no postoperative deaths in the group of patients with an ASA score of 1. Patients with an ASA score of 2 had an observed postoperative mortality rate of 6.2% (11 patients), and in a subsequent ASA score of 3 there were 5 deceased patients (8.3%) were observed. In the highest ASA score of 4, the observed mortality rate increased to 40.0% (2 patients) (Table 5).

By using the HL goodness-of-fit test, no significant difference could be found between the observed and expected frequencies in the ASA classification ($\chi^2 = 1.570$; 1 d.f.; $P = .210$; Table 4). The area under the ROC curve (.635; 95% CI, .51–.76; $P = .055$) did not indicate a discriminatory power. Therefore, the ASA score was not divided into risk categories.

Specification of mortality incidence during the time period

To identify possible differences related to changes in practice over time, the study period (1991–2007) was divided into three 5-year segments. The 90-day mortality rate was not significantly different compared with the overall

Table 4   The 5 risk-adjustment models: calibration and discrimination

Risk-prediction model	HL test (p)	Area under the ROC curve (95% CI)
P-POSSUM	.035	.766 (.67–.86)
O-POSSUM	.529	.756 (.67–.84)
CCI score	.659	.567 (.42–.71)
ACCI score	.270	.684 (.58–.79)
ASA score	.210	.635 (.51–.76)

Table 6   Survival and comorbidity rates in patients during 3 time periods

	Postoperative survivors in periods (%)				Postoperative deceased patients in periods (%)			
	Overall	1991–1996	1997–2002	2003–2007	Overall	1991–1996	1997–2002	2003–2007
90-d mortality	260 (93.5)	81 (94.2)	62 (91.2)	119 (94.4)	18 (6.5)	5 (5.8)	6 (8.8)	7 (5.6)
Comorbidity, yes/no	114/146	24/57	27/35	63/56	13/5	2/3	4/2	7/0
Diabetes mellitus	28 (10.8)	5 (6.2)	8 (12.9)	15 (12.6)	3 (16.7)	1 (20.0)	2 (33.3)	0 (.0)
Hypertension	52 (20.0)	7 (8.6)	8 (12.9)	37 (31.1)	5 (27.8)	0 (.0)	3 (50.0)	2 (28.6)
COPD	33 (12.7)	4 (4.9)	8 (12.9)	21 (17.6)	1 (5.6)	0 (.0)	0 (.0)	1 (14.3)
Angina pectoris	33 (12.7)	11 (13.6)	4 (6.5)	3 (2.5)	3 (16.7)	0 (.0)	0 (.0)	3 (42.9)
Congestive heart failure	3 (1.2)	0 (.0)	0 (.0)	3 (2.5)	0 (.0)	0 (.0)	0 (.0)	0 (.0)
Myocardial infarction	28 (10.8)	10 (12.3)	4 (6.5)	14 (11.8)	3 (16.7)	1 (20.0)	0 (.0)	2 (28.6)
TIA/CVA	14 (5.4)	1 (1.2)	3 (4.8)	10 (8.4)	2 (11.1)	1 (20.0)	0 (.0)	1 (14.3)

COPD= chronic obstructive pulmonary disease; TIA/CVA = transient ischemic attack/cerebrovascular accident.

mortality rate of 6.5%, that is, 5.8% from 1991 to 1996 ($P = .854$), 8.8% from 1997 to 2002 ($P = .396$), and 5.7% from 2003 to 2007 ($P = .721$). However, a significant portion of the patients who died postoperatively had 1 or more severe comorbidities ($P = .018$). Of cardiovascular diseases, which occurred frequently, transient ischemic attack/cerebrovascular accident ($P = .007$) was observed significantly more often from 1991 to 1996, hypertension ($P = .019$) was observed significantly more often between 1997 and 2002, and angina pectoris ($P = .000$) was observed significantly more often between 2003 and 2007 (Table 6). In addition, the predictive power of each model did not differ in these 3 time periods and both POSSUM models had the strongest predictive power in each period.

Comments

Risk stratification in high-risk cancer surgery is pivotal in identifying patients who may benefit from specific perioperative management strategies. Although it is difficult to define risk factors associated with adverse outcome in individual patients, evaluation of postoperative mortality and morbidity is necessary not only for adequate preoperative selection of patients but also for a reliable auditing process comparing outcomes across surgeons and hospitals. In the present study from a single tertiary-care referral center, statistical analyses showed the most accurate individual risk probabilities for O-POSSUM. Overall postoperative mortality was well predicted by the P-POSSUM equation with a low rate of underprediction (n = 1). Therefore, in our cohort the P-POSSUM equation is the most powerful predictor when comparing different cohorts.

There seems to be a contradiction between the overestimated value of postoperative mortality by O-POSSUM and its accurate calibration and discriminatory power for an individual patient. However, predictive accuracy refers to the ability of a model to assign the correct probability of death to patients, whereas discriminatory power refers to the ability of a model to attribute the correct outcomes to patients.[24]

External validation showed varied results regarding prognostic values for these risk-prediction models.[11–13,21–27] Two studies that compared the P-POSSUM and O-POSSUM equations showed a poor HL goodness-of-fit for O-POSSUM, whereas one suggested good predictive power for P-POSSUM.[25,26] Several studies evaluated the O-POSSUM equation and found a variety of results ranging from moderate to good fit.[13,24] Only a few studies were performed to validate the predictive power of the CCI, ACCI, and ASA scores after esophagectomy.[11,12] In a recent study, an association was suggested between a high Charlson score (>2) and mortality[11]; 2 other studies indicated a relationship between mortality and ASA score.[12,21]

These varied results may have had several causes. In the first place, these risk-adjusted models could be interpreted in various ways by investigators. For example, the ASA score is defined by an individual anesthetist at a specific moment and assessments might be influenced by variations in the clinical presentation. Moreover, the ASA score is interobserver-dependent and prone to allocation variation.

A second important difference lies in the definition of mortality. Most of the conducted studies used 30-day mortality as a determinant of surgical outcome, whereas others used in-hospital mortality. In the present study we used the overall postoperative mortality, defined as in-hospital and 90-day mortality. Most of the applied risk prediction models are developed to calculate mortality risk, without any corrections regarding postadmission death within a reasonable period. In applying the 90-day mortality, we included all surgery-related deceased patients. None of these patients died from other circumstances other than the impact of the surgery. Because many patients have a predictably short life span we have to rethink the value of therapeutic strategy if a lot of time would be needed to recuperate from major surgical stress.

Third, because hospital volume appeared to be an important prognostic value,[6] it would be difficult to identify predictive risk factors, particularly in a heterogeneous group. Therefore, we only examined patients who underwent a uniform surgical approach, including a transthoracic esophagectomy with a 2-field lymphadenectomy in a ter-

tiary referral center with a high hospital and surgeon volume. Although still premature, recent literature showed a decrease in severity and frequency of morbidity in patients who underwent a laparoscopic approach, but evidence for a reduced mortality rate has not been established yet.[27,28] None of the patients in this study had a laparoscopic approach and further research is necessary to examine the applicability of these risk-prediction models in such a cohort.

A drawback of this study is in its 16-year time span. A number of factors affecting survival may have evolved over this period such as better patient selection or newer technology, including neoadjuvant chemoradiation and surgical approaches. To counteract the possibility of interfering factors over this period, we divided the time span into 3 almost equal segments. Mortality rates did not differ significantly over this period and statistical analysis indicated the most predictive power for both POSSUM models in each segment. The influence of neoadjuvant chemoradiation in this study was low because there was no mortality in this rather small group of patients (n = 10).

Recently, new risk-adjusted models were developed, including the Rotterdam, Philadelphia, and Munich scores, to compare cohorts, but they do not provide individual risk stratification, as was clearly concluded by Zingg et al.[29]

To date, a reliable individual risk-analysis stratification to guide surgeons and oncologists in the decision making is missing and it should be performed in the context of an overall clinical judgment. With a more appropriate risk-prediction model, we might be able to identify patients with high estimated morbidity and mortality. A careful selection based on such models may be helpful to perform adequate preoperative interventions and reduce the risk of postoperative complications.

Current centralization of this high-risk surgery has led to a relatively low postoperative morbidity rate, and better outcomes have been observed in high-volume centers for moderate- to high-risk patients.[30] Predicting the mortality risk in an individual patient is difficult. The number of events is too few to justify clinical application of any scoring system without further validation by prospective data in a clinical trial setting. To counteract the impossibility of the current models in selecting the individual at-risk patient, we subdivided the most accurate model into a low-, intermediate-, and high-risk category. The benefit of this subdivision for a model is no longer the identification of a rare event, but to identify a group of patients with an increased mortality risk. This subdivision may also be valuable to inform patients participating in clinical studies more accurately and seems to be useful because it is immediately obvious to which risk group a patient belongs. To justify this distribution in clinical practice, more research is necessary to validate this quantification.

Conclusions

Each risk-adjusted model showed a moderate relationship between postoperative mortality and an increased risk

score. We recommend the O-POSSUM for individual risk stratification because it assessed the condition of the patient and the risk of surgery most accurately in this study. In clinical practice we suggest dividing the O-POSSUM score into a low-, intermediate-, and high-risk category, but before general application more research is needed to validate our findings.

References

1. Wu PC, Posner MC. The role of surgery in the management of oesophageal cancer. Lancet Oncol 2003;4:481–8.
2. Kocher HM, Linklater K, Patel S, et al. Epidemiological study of oesophageal and gastric cancer in south-east England. Br J Surg 2001;88:1249–57.
3. Ma JY, Wu Z, Wang Y, et al. Clinicopathologic characteristics of esophagectomy for esophageal carcinoma in elderly patients. World J Gastroenterol 2006;12:1296–9.
4. Moskovitz AH, Rizk NP, Venkatraman E, et al. Mortality increases for octogenarians undergoing esophagogastrectomy for esophageal cancer. Ann Thorac Surg 2006;82:2031–6.
5. Fink U, Stein HJ, Bochtler J, et al. Neoadjuvant therapy for squamous cell esophageal carcinoma. Ann Oncol 1994;5(Suppl 3):17–26.
6. Stein HJ, Siewert JR. Improved prognosis of resected esophageal cancer. World J Surg 2004;28:520–5.
7. Pennathur A, Luketich JD. Resection for esophageal cancer: strategies for optimal management. Ann Thorac Surg 2008;85:S751–6.
8. Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med 2003;349:2241–52.
9. Sauvanet A, Mariette C, Thomas P, et al. Mortality and morbidity after resection for adenocarcinoma of the gastroesophageal junction: predictive factors. J Am Coll Surg 2005;201:253–62.
10. Low DE, Kunz S, Schembre D, et al. Esophagectomy—it's not just about mortality anymore: standardized perioperative clinical pathways improve outcomes in patients with esophageal cancer. J Gastrointest Surg 2007;11:1395–402.
11. Ra J, Paulson EC, Kucharczuk J, et al. Postoperative mortality after esophagectomy for cancer: development of a preoperative risk prediction model. Ann Surg Oncol 2008;15:1577–84.
12. Tekkis PP, McCulloch P, Poloniecki JD, et al. Risk-adjusted prediction of operative mortality in oesophagogastric surgery with O-POSSUM. Br J Surg 2004;91:288–95.
13. Zafirellis KD, Fountoulakis A, Dolan K, et al. Evaluation of POSSUM in patients with oesophageal cancer undergoing resection. Br J Surg 2002;89:1150–5.
14. Liu JF, Watson DI, Devitt PG, et al. Risk factor analysis of postoperative mortality in oesophagectomy. Dis Esophagus 2000;13:130–5.
15. Rice TM, Rusch VW, Ishwaran H, et al. Cancer of the esophagus and esophagastric junction. Cancer 2010.
16. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg 1991;78:355–60.
17. Prytherch DR, Whiteley MS, Higgins B, et al. POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity. Br J Surg 1998;85:1217–20.
18. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Cancer 2010;116:3763–73.
19. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. J Clin Epidemiol 1993;46:1075–9.
20. Charlson M, Szatrowski TP, Peterson J, et al. Validation of a combined comorbidity index. J Clin Epidemiol 1994;47:1245–51.

21. Reid BC, Alberg AJ, Klassen AC, et al. The American Society of Anesthesiologists' class as a comorbidity index in a cohort of head and neck cancer surgical patients. Head Neck 2001;23:985–94.

22. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.

23. Hosmer DW, Hosmer T, Le Cessie S, et al. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 1997;16:965–80.

24. Lai F, Kwan TL, Yuen WC, et al. Evaluation of various POSSUM models for predicting mortality in patients undergoing elective oesophagectomy for carcinoma. Br J Surg 2007;94:1172–8.

25. Nagabhushan JS, Srinath S, Weir F, et al. Comparison of P-POSSUM and O-POSSUM in predicting mortality after oesophagogastric resections. Postgrad Med J 2007;83:355–8.

26. Lagarde SM, Maris AK, de Castro SM, et al. Evaluation of O-POS-SUM in predicting in-hospital mortality after resection for oesophageal cancer. Br J Surg 2007;94:1521–6.

27. Pham TH, Perry KA, Dolan JP, et al. Comparison of perioperative outcomes after combined thoracoscopic-laparoscopic esophagectomy and open Ivor-Lewis esophagectomy. Am J Surg 2010;199:594–8.

28. Sgourakis G, Gockel I, Radtke A, et al. Minimally invasive versus open esophagectomy: meta-analysis of outcomes. Dig Dis Sci 2010; 55:3031–40.

29. Zingg U, Langton C, Addison B, et al. Risk prediction scores for postoperative mortality after esophagectomy: validation of different models. J Gastrointest Surg 2009;13:611–8.

30. Bilimoria KY, Bentrem DJ, Talamonti MS, et al. Risk-based selective referral for cancer surgery: a potential strategy to improve perioperative outcomes. Ann Surg 2010;251:708–16.