A systematic review of the Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity and its Portsmouth modification as predictors of post-operative morbidity and mortality in patients undergoing pancreatic surgery

Haolu Wang, M.D., Tao Chen, M.D., Hui Wang, M.D., Yanyan Song, Ph.D., Xinxing Li, M.D., Jian Wang, M.D., Ph.D.*

Department of General Surgery, Renji Hospital, Shanghai Jiao Tong University School of Medicine, 1630 S. Dongfang Road, Shanghai, 200127, China; Department of Biostatistics, Institutes of Medical Sciences, Shanghai Jiao Tong University School of Medicine, 280 S. Chongqing Road, Shanghai, 200025, China

**KEYWORDS:** Morbidity; Mortality; Pancreatic surgery; POSSUM; P-POSSUM; Systematic review

**Abstract**

**BACKGROUND:** The Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) model and its Portsmouth modification (P-POSSUM) are used extensively to predict postoperative mortality and morbidity in general surgery. The aim of this study was to undertake the first systematic review of the predictive value of these models in patients undergoing pancreatic surgery.

**METHODS:** Eligible articles were identified by searches of electronic databases for those published from 1991 to 2012. Two independent reviewers assessed each study against inclusion and exclusion criteria. All data were specific to pancreatic surgery. Predictive value of morbidity and mortality were assessed by calculating observed/expected ratios.

**RESULTS:** Nine studies were included in the final review. The morbidity analysis included 8 studies (1,734 patients) of POSSUM with a weighted observed/expected ratio of .85. The mortality analysis included 5 studies (936 patients) of POSSUM and 4 studies (716 patients) of P-POSSUM. Weighted observed/expected ratios for mortality were .35 for POSSUM and 1.39 for P-POSSUM.

**CONCLUSIONS:** POSSUM overpredicted postoperative morbidity in patients undergoing pancreatic surgery. Both POSSUM and P-POSSUM failed to offer significant predictive value for mortality in pancreatic surgery, and more data collection in large populations undergoing pancreatic surgery are needed.

© 2013 Elsevier Inc. All rights reserved.

Although operative technique and postoperative care have developed, pancreatic surgery remains associated with high morbidity and mortality. Postoperative complications such as primarily pancreatic fistula, hemorrhage, abscess, and delayed gastric emptying still occur with 30% to 60% frequency, resulting in mortality of 1% to 5%. Many studies suggest that various prognostic factors (eg, age, disease severity, and type of operation) influence surgical...
outcomes. Thus, surgical audit on the basis of these identified risk factors can be used in early detecting postoperative complications and allowing optimized treatment planning to improve individual prognosis.

In a number of audit systems devised over the past 2 decades, the Physiologic and Operative Severity Score for the enUmeration of Mortality and morbidity (POSSUM) model of Copeland et al was recognized as most effective for general surgery. This model, using scores relating to 12 physiologic and 6 operative variables, was developed to predict 30-day mortality and morbidity postoperatively. However, POSSUM was then found to overpredict postoperative mortality, particularly in patients undergoing minor procedures. The Portsmouth modification of POSSUM (P-POSSUM) was developed by Whiteley et al using the same physiologic and operative variables but a linear method rather than an exponential method of regression analysis to predict mortality. Recently, the precision of predicting postoperative morbidity and mortality using the POSSUM and P-POSSUM scoring systems has been reviewed in gastroesophageal and colorectal surgery.

The application of the predictive POSSUM and P-POSSUM models to pancreatic surgery has generated conflicting results. The aim of the present study was to undertake a systematic review of the value of POSSUM and P-POSSUM in predicting postoperative morbidity and mortality in patients undergoing pancreatic surgery.

Methods

Search strategy

This systematic review was undertaken according to a predefined simplified protocol. PubMed, Embase, the Cochrane Controlled Trials Register, the Science Citation Index, and the Chinese Biomedical Database were used for systematic literature searches. Eligibility was restricted to studies published between 1991 (when the original POSSUM model was published) and 2012 with abstracts or full text available. As search terms, we used “POSSUM” combined with each of the following: “pancreas,” “pancreatic,” “pancreatitis,” “pancreatectomy,” and “pancreatoduodenectomy.” A comprehensive search of the reference lists of all review articles and original studies retrieved by this method was performed to identify additional reports. No language restrictions were used.

Criteria for review

The original articles of those studies that examined POSSUM scoring system (POSSUM or P-POSSUM) prospectively or retrospectively in pancreatic surgery were included in our study. Studies in abstract form and reports presented at meetings without full publication were also included in the analysis. Studies were not included if they did not provide details on observed and expected events of POSSUM or P-POSSUM.

Review procedure

Titles and abstracts were studied to assess inclusion criteria. All articles were examined independently for eligibility by 2 reviewers (Haolu Wang, T.C.). Disagreements were resolved by consulting a third reviewer (J.W.).

Data analysis

Postoperative morbidity and mortality were assessed by 30-day or in-hospital rates. The ratio of observed to expected events (O/E ratio) of each POSSUM model was recorded. The weighted O/E ratio was calculated and plotted dependent on the percentage of each sample size (the number of patients in each study) in the total using SAS (SAS Institute Inc, Cary, NC). An O/E ratio of 1 confers 100% predictive accuracy, a ratio <1 implies model overprediction of events, and a ratio >1 implies model underprediction of events.

Results

Description of the studies

The bibliographic search yielded a total of 126 studies, and full text or abstracts were obtained for 17 studies. Eight of these studies did not meet the inclusion criteria: 1 did not contain data specific to pancreatic surgery, 4 did not provide details on observed and expected events of POSSUM or P-POSSUM, 1 was performed to evaluate bypass procedures in patients with unresectable pancreatic cancer, and 2 referred to the same data set. Nine independent studies were thus included in the final review. The selection process and exclusion criteria applied are summarized in Fig. 1.

Studies included in the morbidity analysis

Eight studies, comprising data on 1,734 patients, reported the accuracy of POSSUM in predicting postoperative morbidity after pancreatic surgery. The patient and operative characteristics of these studies are shown in Table 1. One study reported morbidity on the basis of the original POSSUM definitions, and 1 study referred to the International Study Group of Pancreatic Surgery definitions of complications after pancreatectomy. Three studies identified morbidity according to the Clavien complication scheme, and 3 studies used an arbitrary list of complications. There were also no significant differences in the operations performed. Six studies reported data on pancreaticoduodenectomy, while 2 others reported on pancreatic resection including pancreaticoduodenectomy.
Studies included in the mortality analysis

Five studies, comprising data on 936 patients, reported the accuracy of POSSUM in predicting postoperative mortality after pancreatic surgery, and 4 studies, with 716 patients, reported the accuracy of P-POSSUM. These studies are summarized in Table 2 and Table 3. Three studies of POSSUM\(^{19,20,23}\) and 2 studies of P-POSSUM\(^{19,23}\) reported death rates as in-hospital mortality, while 1 study of POSSUM\(^{24}\) and 1 study of P-POSSUM\(^{26}\) reported 30-day mortality.

There were also no significant differences in the operations performed. Three studies of POSSUM\(^{20,23,24}\) and 2 studies of P-POSSUM\(^{23,26}\) reported data in patients undergoing pancreaticoduodenectomy, while 2 others\(^{18,19}\) reported on pancreatic resection including pancreaticoduodenectomy.

Observed/expected ratios for morbidity

The O/E ratios of postoperative morbidity for POSSUM in the pancreatic resection studies are shown in Table 4, and those for the pancreaticoduodenectomy studies are shown in Table 5. The observed morbidity of studies of POSSUM in pancreatic resection ranged from 26.0% to 54.1%, while the predicted mortality ranged from 37.8% to 70%. The weighted O/E ratio for morbidity for POSSUM was .85 (range, .50–.96) in pancreatic resection and .85 (range, .50–.96) in pancreaticoduodenectomy.

Observed/expected ratios for mortality

The O/E ratios for postoperative mortality for POSSUM and P-POSSUM in the pancreatic resection studies are shown in Table 4, and those in the pancreaticoduodenectomy studies are shown in Table 5. The observed mortality of studies of POSSUM in pancreatic resection ranged from 1.2% to 5.6%, while the predicted mortality ranged from 6.2% to 27.8%. There was great heterogeneity regarding the observed mortality of the pancreatic resection studies of P-POSSUM (range, 1.2%–7.8%). The predicted mortality for P-POSSUM in pancreatic resection ranged from 2.3% to 6.5%. The weighted O/E ratios for mortality for POSSUM and P-POSSUM in pancreatic resection were .35 (range, .07–.67) and 1.39 (range, .19–3.4), respectively. The weighted O/E ratios for mortality for POSSUM and P-POSSUM in pancreaticoduodenectomy were .55 (range, .20–.67) and 2.93 (range, .67–3.4), respectively.
Comments

The present study is, to our knowledge, the first systematic review of POSSUM as a predictor of postoperative morbidity and mortality in patients undergoing pancreatic surgery. Until now, no clear results have existed regarding the use of POSSUM models to predict postoperative morbidity and mortality in pancreatic surgery. Khan et al

were the first to evaluate POSSUM for pancreatic surgery, and they found that the model overestimated postoperative morbidity and mortality. However, that study was limited by the small number of patients (n = 50). A larger and more recent study suggested that POSSUM was a good predictor of morbidity, but the model had a lack of fit for mortality.

The results of the present review of 1,734 pancreatic operations suggest that POSSUM overpredicted postoperative morbidity (O/E ratio, .85) in patients undergoing pancreatic surgery. Both POSSUM and P-POSSUM for postoperative mortality prediction had clear limitations. POSSUM overpredicted postoperative mortality for

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Studies of POSSUM for postoperative morbidity in patients undergoing pancreatic surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>Year (country)</td>
</tr>
<tr>
<td>Khan et al23</td>
<td>2003 (United Kingdom)</td>
</tr>
<tr>
<td>Zhang et al24</td>
<td>2006 (China)</td>
</tr>
<tr>
<td>Pratt et al19</td>
<td>2007 (United States)</td>
</tr>
<tr>
<td>Zhang et al20</td>
<td>2009 (China)</td>
</tr>
<tr>
<td>de Castro et al17</td>
<td>2009 (The Netherlands)</td>
</tr>
<tr>
<td>Knight et al18</td>
<td>2010 (United Kingdom)</td>
</tr>
<tr>
<td>Debinska et al25</td>
<td>2011 (Poland)</td>
</tr>
<tr>
<td>Gallacher et al21</td>
<td>2011 (United Kingdom)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Studies of POSSUM for postoperative mortality in patients undergoing pancreatic surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>Year (country)</td>
</tr>
<tr>
<td>Khan et al23</td>
<td>2003 (United Kingdom)</td>
</tr>
<tr>
<td>Zhang et al24</td>
<td>2006 (China)</td>
</tr>
<tr>
<td>Pratt et al19</td>
<td>2007 (United States)</td>
</tr>
<tr>
<td>Zhang et al20</td>
<td>2009 (China)</td>
</tr>
<tr>
<td>Knight et al18</td>
<td>2010 (United Kingdom)</td>
</tr>
</tbody>
</table>

O/E = observed/expected; PD = pancreaticoduodenectomy; POSSUM = Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity.

*Equivalent in patients aged >65 years (n = 68; P > .05, χ² test), overestimates risk in patients <65 years (n = 128; P < .05; χ² test).

†Chi-square test or Fisher’s exact test.

‡Goodness-of-fit analysis.
936 pancreatic operations significantly (O/E ratio, .35) and P-POSSUM underpredicted mortality for 716 pancreatic operations (O/E ratio, 1.39). The main reason for the lack of predictive value of POSSUM models may be that POSSUM was introduced in 1991, but the application of POSSUM and P-POSSUM in predicting outcomes after pancreatic surgery has been reported only since 2003. During those 12 years, surgical techniques improved, and better perioperative care was developed at more specialized centers. Risk-scoring systems developed in the late 1980s may no longer be applicable. Also, POSSUM was originally developed using data from general surgical patients. It has been reported that POSSUM might overpredict morbidity and mortality in special general surgical procedures, especially in procedures involving a high degree of difficulty.27 In patients undergoing pancreatic operations, surgery has a high operative severity score and therefore results in a higher risk prediction. Compared with POSSUM, specialty-specific models, including V-POSSUM for vascular surgery, O-POSSUM for gastrointestinal surgery, and CR-POSSUM for colorectal surgery, could improve the reliability of the prediction of postoperative morbidity and mortality.5,6 Finally, another possible explanation why POSSUM did not produce reliable prognoses in our review might be that the risk factors of complications in pancreatic surgery differ from those in gastrointestinal, colorectal, and other general surgical procedures. Serum bilirubin, for example, is not scored in POSSUM but is known as an important prognostic factor of hepatobiliary-pancreatic cancer.5,27,28 Therefore, we suggest that a pancreatic surgery–specific revision of POSSUM is needed in the future. However, the reasons for the lack of value of P-POSSUM are not clear. Significant different findings were observed in included studies (O/E ratio range, .19–3.4).

Despite these limitations, there is still a role for POSSUM as an audit tool in pancreatic surgery. Because of the overprediction of postoperative morbidity and mortality, individual POSSUM scores should not preclude pancreatic resection in clinic practice but could help surgeons alter expectations of postoperative outcomes.19,20 POSSUM models could be used to help guide management scenarios in pancreatic surgical practice. For example, patients with high physiologic POSSUM scores can be provided additional preoperative interventions such as enteral tube feedings, hyperalimentation, antibiotics, and biliary stenting, to improve physiologic parameters. At some high-volume centers, this is often indicated, particularly when patients present with malignant obstructive jaundice, comorbid cardiac or respiratory illness, diabetes, or malnutrition.19

There are several reasons for the different findings from different studies. First, given that postoperative deaths are considerably less common than postoperative complications...
in patients undergoing pancreatic surgery, there would be a wider variation in reported O/E ratios of mortality than that of morbidity. Second, a number of studies identified and included in the present review involved pancreaticoduodenectomy alone instead of pancreatic surgery. It has been observed in some studies that because the potential for morbidity and mortality is greater after pancreaticoduodenectomy, POSSUM and P-POSSUM demonstrated more predictive value in this surgical procedure.19 Thus, we evaluated the corresponding results if only the studies on pancreaticoduodenectomy were included. POSSUM indeed had more accuracy in predicting postoperative mortality for pancreaticoduodenectomy (O/E ratio, .55) than for pancreatic surgery (O/E ratio, .35). However, additional value was found neither for POSSUM in predicting morbidity after pancreaticoduodenectomy (O/E ratio, .85) nor for P-POSSUM in predicting mortality (O/E ratio, 2.93). Therefore, it remains unclear whether the type of resection influences the validity of the scores. Third, some studies were heterogeneous in case mix. Studies from Europe cast serious doubt on POSSUM17,18,21,23,25 in predicting postoperative morbidity for pancreatic surgery, but reports from the United States19 and China20 supported its use. A possible reason for these varied results of POSSUM could be the use of different definitions for what constitutes a postoperative complication. For example, the International Study Group on Pancreatic Fistula found that different definitions of pancreatic leakage exist, and the reported range from 2% to 50% underscores this variation.17,29 Another reason is that specific management or technical approaches in some studies might contribute to different outcomes. It is notable that the 1 study that strongly supported the use of POSSUM was from a high-volume institution at which all 326 pancreatic resections were performed by 2 fellowship-trained hepatopancreatobiliary surgeons.19 Therefore, the inclusion of heterogeneous studies might be a confounding factor in the present systematic review. This is also a limitation of the present systemic review. We should exert caution when interpreting the weighted O/E ratio for morbidity and mortality, because the original studies were not powered to homogeneously and specifically evaluate this outcome. Fortunately, all studies provided comparable mortality and morbidity data, and most procedures performed in the 2 studies on pancreatic surgery18,19 were also pancreaticoduodenectomy, allowing a comparatively fair evaluation of these outcomes.

Considering all available evidence together, POSSUM overpredicted postoperative morbidity in patients undergoing pancreatic surgery. Both POSSUM and P-POSSUM failed to offer significant predictive value for mortality in pancreatic surgery. This has implications for clinical practice because there appears to be insufficient evidence to promote their use in pancreatic surgery. More trials are needed to adequately evaluate POSSUM and P-POSSUM in predicting postoperative mortality for pancreatic surgery.

**Acknowledgments**

We thank Dr Hao Fang, Southeast University School of Information Science and Engineering, for his technical assistance in this study.

**References**
