Validation of the Surgical Outcome Risk Tool (SORT) in patients with pancreatic cancer undergoing surgery

Dear Editor,

Pancreatic resection is currently accepted as the mainstay of the multimodal treatment strategy for resectable and borderline pancreatic cancer. In this context, the evaluation of patients' clinical status, along with the risk of perioperative morbidity and mortality for this type of major surgery is crucial to support the shared decision-making process, along with enhancing the oncologic treatment strategy, counseling, and outcome. The Surgical Outcome Risk Tool (SORT) was developed following the 2011 National Confidential Enquiry into Patient Outcome and Death (NCEPOD) report, in order to provide enhanced identification of high-risk surgical patients.[1] To achieve this goal, SORT employs only six variables, designed to predict a patient's probability of 30-day postoperative mortality. Currently, it has been compared favorably with other risk stratification tools and has been externally validated in patients undergoing hip fracture and liver surgery.[2,3] However, it has not been validated for a pancreatic cancer surgical population. The purpose of the present study was to validate the SORT model in Greek adult patients undergoing surgery for pancreatic cancer. We also compared SORT with two additional risk stratification tools, the Physiology and Operative Severity Score for the enumeration of Mortality and Morbidity (POSSUM), and the Portsmouth POSSUM (P-POSSUM).

Data were obtained from a prospectively maintained database of consecutive patients undergoing surgery for pancreatic cancer between January 1st, 2017 and December 31st, 2019 and ethical approval was obtained by the Scientific Committee of the University Hospital of Larissa, Greece (Protocol number: 50271/30-10-19). All the procedures were performed by the same surgical team leaded by the senior author (DZ). No imputation methods were used regarding missing data. We assessed the discrimination (i.e., the ability to separate those who died from those who did not die) and calibration (i.e., the ability to predict mortality rates in agreement with actual observed mortality rates) of the SORT model. Discrimination was assessed by generating Receiver-Operating Characteristic (ROC) curves and by calculating the Area Under the ROC Curve (AUC). The AUC was determined by calculating the 95% Confidence Intervals and compared using nonparametric paired tests, as described by DeLong et al.[4] We defined as poor, fair, and excellent model discrimination the AUC of < 0.70, 0.70–0.79 and 0.80–1.00, respectively. The calibration regarding each model was evaluated by estimating the predicted mortality (expected) and then comparing with the true mortality (observed). The observed/expected ratio of 1 represents perfect accuracy, a ratio < 1 indicates overprediction of mortality rate, and a ratio > 1 indicates underestimation. Calibration was further evaluated using the Hosmer-Lemeshow (H-L) goodness of fit test, defining a lack of fit as a p-value ≤ 0.05.[5] Finally, Chi-squared testing was used to compare the observed and expected outcome of all patients. All data were analyzed using Microsoft® Excel 16.36 (Microsoft, Redmond, Washington, USA) and Prism® Graphpad 8.4.2 for MacOS (GraphPad Software, San Diego, CA).

Table 1 Patient baseline characteristics.

Demographics	Number of patients, n = 50
Female, n (%)	21 (42)
Mean age, years (SD)	66.6 (11)
ASAphysical status, n (%)	
I	10 (20)
II	23 (46)
III	15 (30)
IV	2 (4)
Operation priority, n (%)	
Elective	50 (100)
Acute	0 (0)
Surgical operations, n (%)	
Pancreaticoduodenectomy	42 (84)
Total pancreatectomy	1 (2)
Distal pancreatectomy	7 (14)
30-day mortality, n (%)	3 (6)

ASA, American Society of Anesthesiologists.

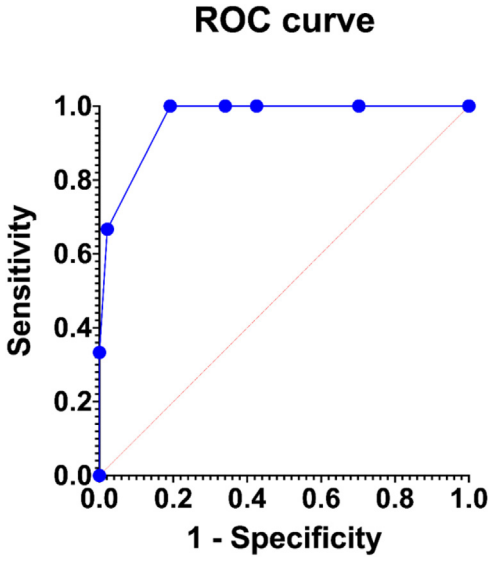

Figure 1 The Receiver Operating Characteristics (ROC) Curve demonstrating the discrimination level of the Surgical Outcome Risk Tool (SORT) in patients with pancreatic cancer undergoing surgery.

Fifty patients with pancreatic cancer were incorporated in the present analysis (Table 1), with a mean age of 66.7 years. The mean length of hospital stay was 17.52 (±7.29) days and the mean length of stay in the intensive care unit was 0.98 (±0.42) days. In the current study we reported a 30-day mortality rate of 6% (3 patients). SORT was associated with an excellent discrimination level (AUC = 0.96 [95% CI: 0.89–1.00]; $p = 0.008$). The ROC curve is demonstrated in Figure 1. SORT also demonstrated a significantly low H-L value (H-L = 0.02; $p > 0.99$), thus passing the goodness of fit test. Nonetheless, it underestimated the mortality rate (O:E = 1.5). POSSUM demonstrated a lower discrimina-

tion level (AUC = 0.89 [95% CI: 0.70–1.00]; $p = 0.026$) and a higher H-L value (H-L = 1.77; $p = 0.99$). It also underestimated mortality (O:E = 1.5). P-POSSUM was also associated with an excellent discrimination level (AUC = 0.95 [95% CI: 0.87–1.00]; $p = 0.010$), but lower than SORT, while underestimating the mortality rate at a higher level compared with SORT (O:E = 3). In addition, P-POSSUM was associated with a higher H-L value (H-L = 1.58; $p = 0.99$) in comparison to SORT.

There are certain limitations to the present study. In fact, the design of the study was retrospective, and the study population was small. Nonetheless, this is the first evidence regarding the validity of SORT in patients with pancreatic cancer undergoing surgery. In addition, we demonstrated that SORT is associated with excellent discrimination and an appropriate level of calibration in predicting postoperative mortality. Furthermore, our outcomes suggest the superiority of SORT compared with POSSUM and P-POSSUM. Future studies should further assess SORT in a greater study population of patients with pancreatic cancer undergoing surgery, with a greater follow-up, along with comparing it with other risk assessment tools.

Ethical approval

Ethical approval was obtained by the Scientific Committee of the University Hospital of Larissa (Protocol nº 50271/30-10-19).

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Protopapa KL, Simpson JC, Smith NC, et al. Development, and validation of the Surgical Outcome Risk Tool (SORT). Br J Surg. 2014;101:1774–83.
2. Metz CE, Herman BA, Roe CA. Statistical comparison of two ROC-curve estimates obtained from partially paired datasets. Med Dec Making. 1998;18:110–21.
3. Wong GTC, Ang WC, Wong TCL, et al. Surgical Outcome Risk Tool (SORT) validation in hepatectomy. Anaesthesia. 2017;72:1287–9.
4. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44:837–45.
5. 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.

Dimitrios E. Magouliotis[a,b], Athina Samara[b], Maria P. Fergadi[b], Dimitrios Symeonidis[b], Dimitris Zacharoulis[b,*]

[a] *UCL, Faculty of Medical Sciences, Division of Surgery and Interventional Science, London, UK*
[b] *University of Thessaly, Department of Surgery, Biopolis, Larissa, Greece*

* Corresponding author.
E-mail: zacharoulis@uth.gr (D. Zacharoulis).

Received 15 May 2020; accepted 31 October 2020

https://doi.org/10.1016/j.bjane.2020.10.018

Airway management in obese patients

Dear Editor,

We read with interest the article of Turna et al. [1] on their randomized trial of performance of the Airtraq videolaryngoscope versus the intubating laryngeal mask airway (ILMA) in obese patients. There are several aspects to the study we believe necessary to consider.

Airway management in obese patients is a challenging issue associated with a high incidence of complications. The accumulation of adipose tissue causes several changes in airway anatomy and respiratory function. Thus, obesity is associated with, among others, decreased pharyngeal area, obstructive sleep apnea, restrictions in neck flexion, narrow jaw opening, enlarged tongue, reduction in functional residual capacity and alveolar oxygen reserve, and increase in O_2 consumption. Therefore, obese patients are at increased risk of difficult mask ventilation, difficult tracheal intubation, and hypoxemia during the process of securing the airway, even after short periods of apnea. The core recommendations of the recent guidelines focus on limiting the duration and number of attempts at tracheal intubation in order to achieve early atraumatic intubation, the philosophy on which the vortex approach is based. Accordingly, an undue number of attempts to test a device is not justified. Thereby, it was published in 2016 a useful consensus on airway research ethics that every researcher should take into account. [2] It recommends limiting to a maximum of two failed attempts before following the usual progression in the airway management algorithm and restricting the inclusion of patients to ASA I and II to minimize harm.

Likewise, direct laryngoscopy could not be the most suitable rescue method after the unsuccessful use of a videolaryngoscopy or an ILMA given that its probability of success can be lower in this situation. Perhaps, it would have been more appropriate to use the other device under study as a backup plan. In addition, any blind technique should be avoided due to the significant failure rate, the frequent need for repeated attempts, and the potential for airway trauma, which can result in deterioration of ventilation. [3] Therefore, fiberoptic intubation through the ILMA is the method recommended.

On the other hand, testing a laryngeal video mask as the Totaltrack VLM (Medcomflow S.A., Barcelona, Spain) instead