




Validation of the Surgical Outcome Risk Tool (SORT) for Predicting Postoperative Mortality in Colorectal Cancer Patients Undergoing Surgery and Subgroup Analysis

Dimitrios E. Magouliotis^{1,2} · David Walker¹ · Ioannis Baloyiannis² · Maria P. Fergadi³ · Ioannis Mamaloudis² · Georgios Chasiotis² · George A. Tzovaras² 

Accepted: 30 January 2021 / Published online: 18 February 2021
© Société Internationale de Chirurgie 2021

Abstract

Background The accurate evaluation of perioperative risk is crucial to facilitate the shared decision-making process. Surgical outcome risk tool (SORT) has been developed to provide enhanced and more feasible identification of high-risk surgical patients. Nonetheless, SORT has not been validated for patients with colorectal cancer undergoing surgery. Our aim was to determine whether SORT can accurately predict mortality after surgery for colorectal cancer and to compare it with traditional risk models.

Method 526 patients undergoing surgery performed by a colorectal surgical team in a single Greek tertiary hospital (2011–2019) were included. Five risk models were evaluated: (1) SORT, (2) Physiology and Operative Severity Score for the enumeration of Mortality and Morbidity (POSSUM), (3) Portsmouth POSSUM (P-POSSUM), (4) Colorectal POSSUM (CR-POSSUM), and (5) the Association of Great Britain and Ireland (ACPGBI) score. Model accuracy was assessed by observed to expected (*O:E*) ratios, and area under Receiver Operating Characteristic curve (AUC).

Results Ten patients (1.9%) died within 30 days of surgery. SORT was associated with an excellent level of discrimination [AUC:0.81 (95% CI:0.68–0.94); $p = 0.001$] and provided the best performing calibration of all models in the entire dataset analysis (H–L:2.82; $p = 0.83$). Nonetheless, SORT underestimated mortality. SORT model demonstrated excellent discrimination and calibration predicting perioperative mortality in patients undergoing (1) open surgery, (2) emergency/acute surgery, and (3) in cases with colon-located cancer.

Conclusion SORT is an easily adopted risk-assessment tool, associated with enhanced accuracy, that could be implemented in the perioperative pathway of patients undergoing surgery for colorectal cancer.

Supplementary information The online version contains supplementary material available at (<https://doi.org/10.1007/s00268-021-06006-6>).

✉ George A. Tzovaras
gtzovaras@med.uth.gr

Dimitrios E. Magouliotis
dimitrios.magouliotis.18@ucl.ac.uk

David Walker
d.walker@ucl.ac.uk

Ioannis Baloyiannis
balioan@hotmail.com

Maria P. Fergadi
mfergadi@uth.gr

Ioannis Mamaloudis
imamaloudis@yahoo.gr

Georgios Chasiotis
geochas92@gmail.com

¹ Division of Surgery and Interventional Science, Faculty of Medical Sciences, UCL, London, UK

² Department of Surgery, University of Thessaly, Biopolis, 41110 Larissa, Greece

³ Department of Radiology, University of Thessaly, Biopolis, Larissa, Greece

Introduction

Colorectal cancer (CRC) is a common cancer and one of the leading cancer-related causes of death worldwide, accounting for approximately 35,000 new cases and over 16,000 deaths per year [1], with the majority of patients undergoing potentially curative surgery. In this context, the accurate evaluation of perioperative risk is crucial in order to facilitate the shared decision-making and informed consent processes regarding patients undergoing surgery and to enhance the quality of clinical practice during the perioperative pathway. Moreover, the implementation of an accurate risk stratification tool enables the actual comparison of surgical outcomes among different healthcare providers for either service evaluation or clinical audit. Certain risk stratification tools have been introduced into clinical practice [2]. Risk stratification tools may be subdivided into risk scores and risk prediction models. Both are usually developed using multivariate analysis of risk factors for a specific outcome [2]. Despite, the increasing interest in more advanced risk prediction methods, the risk stratification models remain the most easily accessible choice for this purpose. Nonetheless, they are not commonly used in our daily clinical practice, possibly due to poor awareness amongst clinicians, as well as, concerns regarding their complexity and accuracy [3].

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 in a more feasible manner [3]. To achieve this goal, the SORT model uses only six routinely collected data items, designed to predict patient's probability of 30-day postoperative mortality. Currently, it has been compared favorably with other previously validated risk stratification tools, such as the ASA physical status (ASA PS) grade, and has been externally validated in groups of patients undergoing hip fracture surgery [4], along with hepatectomy [5]. In both groups [4, 5] SORT was associated with an acceptable discrimination level [AUC: 0.70 (0.66–0.74) and 0.822 (0.728–0.916), respectively], but showed low calibration traits. However, it has not been validated for a colorectal cancer surgical population, yet. The purpose of the present study was to validate the SORT model in Greek adult patients undergoing surgery for colorectal cancer, along with performing subgroup sensitivity analysis.

Methods

Data extraction

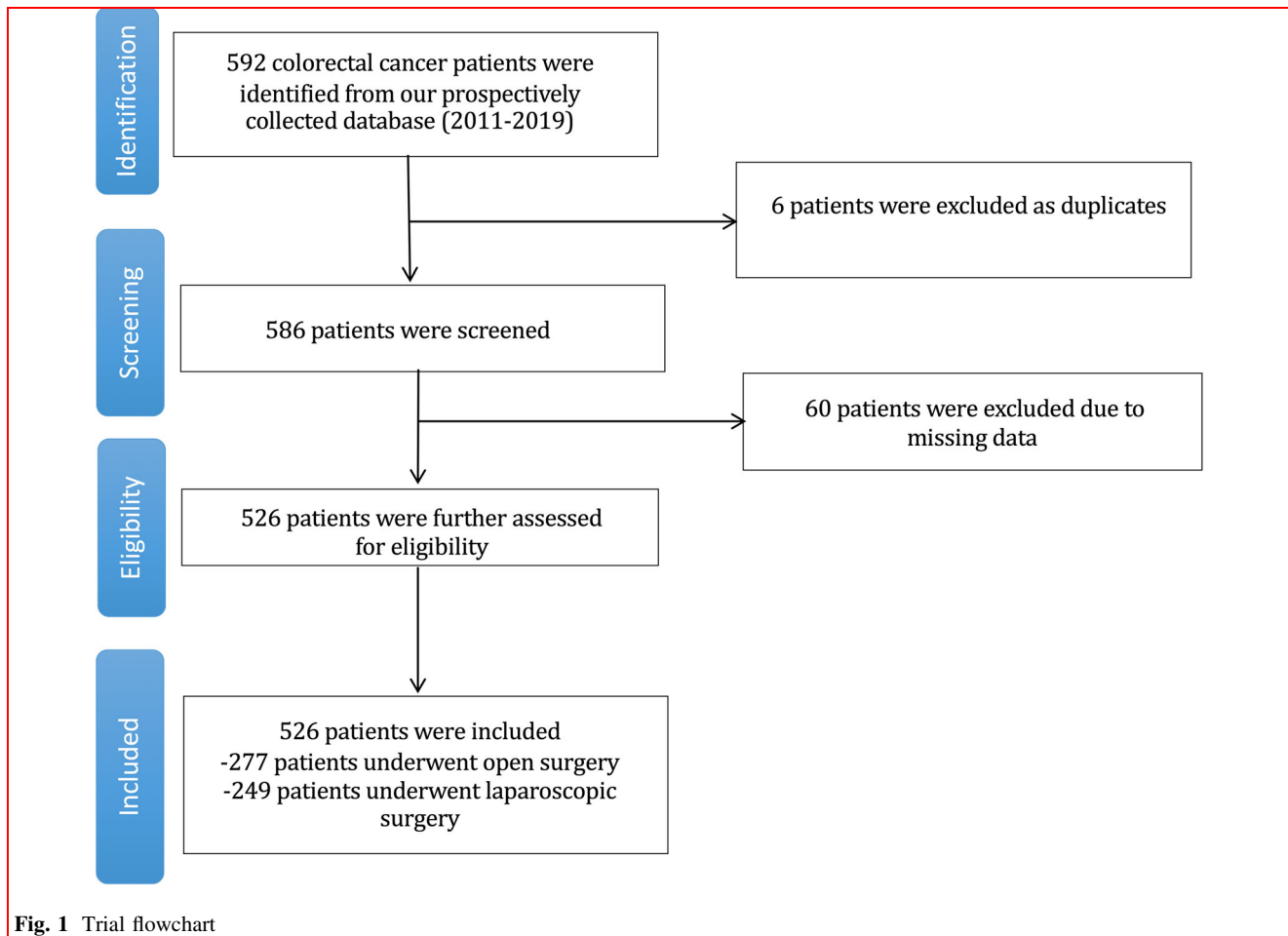
The present study was conducted under the protocol agreed by all authors. Data were obtained from a prospectively maintained database of consecutive patients undergoing surgery for colorectal cancer between January 1st 2011 and December 31st 2019. All the procedures were performed by the same surgical team leading by the senior author (GT) at the Department of Surgery, University Hospital of Larissa, Greece. The choice between open or laparoscopic approach was depended either upon patient's preference (many people in Greece believe that laparoscopic colectomy is still somewhat "experimental") or logistics (time/theatre space, availability of disposables, etc.). Ethical approval was obtained by the Scientific Committee of the University Hospital of Larissa (Protocol number: 33606/16-07-19). Informed consent was not necessary due to the retrospective nature of the present study.

Data on age, gender, surgical approach (laparoscopic or open), American Society of Anesthesiology (ASA) grade, operative priority, surgical severity, malignancy status, staging and type of procedure were prospectively collected. Mortality was defined as any death that occurred during the first 30 days or within the index hospital admission if longer than 30 days. The predicted risk of mortality was determined using SORT model. In addition, we calculated predicted mortality using POSSUM, P-POSSUM, CR-POSSUM, and ACPGBI CRC for all patients. Patients with incomplete data were excluded from analysis.

A separate sensitivity analysis was performed to determine the accuracy of each model to predict perioperative mortality in patients undergoing surgery for colorectal cancer based on (1) procedure-related variables: surgical approach (laparoscopic vs open), operative priority (elective vs acute), (2) a cancer-related variable: cancer site (colon vs rectum), and (3) a patient-related variable: age (≥ 80 vs < 80). We used these variables because they might affect the perioperative mortality.

Endpoints

The primary endpoint of the study was defined as the validation of SORT model in Greek adult patients undergoing surgery for colorectal cancer. Secondary endpoints included (1) the comparison of SORT with POSSUM, P-POSSUM, CR-POSSUM, and ACPGBI CRC models regarding their accuracy in predicting perioperative mortality and (2) subgroup sensitivity analysis.



Statistical analysis

The calculation of the SORT score was performed by employing the method and the web-based calculator developed by Protopappa et al. [3]. The SORT model implements the following variables: ASA physical status (PS), operative priority (elective, urgent, immediate), surgical specialties (gastrointestinal, thoracic, or vascular surgery), surgical severity (major / complex); malignancy status, and age (65–79 or ≥ 80 years). POSSUM, P-POSSUM, CR-POSSUM, and ACPGIBI scores were calculated using the method described by Copland et al. [6], Prytherch et al. [7], Tekkis et al. [8], and Ferjani et al. [9], respectively.

We assessed the discrimination (i.e., the ability to separate those who did 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. [10]. We defined as poor, fair and excellent model discrimination the AUC of < 0.70 , $0.70–0.79$ and $0.80–1.00$, respectively [10].

The calibration regarding each model was calculated by estimating the predicted mortality (expected) and then compared 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 of > 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 [11]. Finally, Chi-squared testing was used to compare the observed and expected outcome of all patients.

All data were analyzed using Microsoft® Excel 2019 (Microsoft, Redmond, Washington, USA) and Prism® Graphpad 8.4 for Mac (GraphPad Software, San Diego, CA).

Table 1 Patient baseline characteristics

Demographics	Number of patients, <i>n</i> = 526
Female, <i>n</i> (%)	193 (36.7)
Mean age, years (SD)	69.75 (10.46)
Age ≥ 80 (%)	90 (17.1)
Age < 80 (%)	436 (82.9)
ASA Class, <i>n</i> (%)	
I	139 (26.4)
II	279 (53.0)
III	104 (19.8)
IV	4 (0.8)
Stage, <i>n</i> (%)	
I	217 (41.3)
II	148 (28.1)
III	120 (22.8)
IV	41 (7.8)
Operative priority	
Elective	464 (88.2)
Acute	62 (11.8)
Cancer site	
Colon	343 (65.2)
Rectum	183 (34.8)
Surgical approach	
Open	277 (52.7)
Laparoscopy	249 (47.3)
30-day mortality	10 (1.9)

ASA American society of anesthesiologists

Results

Baseline characteristics

We report our outcomes according to The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [12]. The Trial Flow regarding data extraction strategy is presented in Fig. 1. A total of 526

patients were included. Patients' baseline characteristics are shown in Table 1. One hundred ninety-three patients (36.7%) were females with a mean age of 69.75 (Standard Deviation—SD: 10.46) years. The majority of patients presented stage I/II cancer (69.4%) and underwent an elective procedure (88.2%). The tumor was located in the colon in 343 (65.2%) patients. A total of 277 (52.7%) patients underwent open surgery and 249 (47.3%) laparoscopy. Finally, the overall mortality rate was 1.9%.

Performance of SORT model in the entire dataset

As demonstrated in Table 2 and Fig. 2, SORT was associated with a high level of discrimination in the total study population [AUC: 0.81 (95% CI: 0.68–0.94); $p = 0.001$]. Furthermore, SORT presented the lowest Hosmer–Lemeshow value (H–L: 2.82; $p = 0.83$), thus providing the best performing calibration of all models in the entire dataset analysis. Nonetheless, SORT underestimated mortality determined by observed/expected ratios of > 1.

Comparison of SORT with other mortality prediction models in the entire dataset

P-POSSUM was also associated with a high discrimination level [AUC:0.85 (95% CI: 0.76–0.94); $p = 0.005$]. POSSUM, CR-POSSUM and ACPGBI CRC models demonstrated a fair discrimination level (Table 2, Fig. 2). While SORT presented the best performing calibration, P-POSSUM demonstrated the worst performing calibration traits (H–L: 8.29; $p = 0.41$) (Table 2, Fig. 2).

Performance of mortality prediction models in subgroups

Subgroup analysis outcomes are provided in Table 3. SORT model demonstrated high discrimination predicting perioperative mortality in patients undergoing (1) open surgery, (2) emergency/acute surgery and (3) in cases with

Table 2 Discrimination and calibration of the studied scores for predicting mortality in colorectal patients

Scoring systems	<i>O</i>	<i>E</i>	<i>O:E</i>	Discrimination		Calibration	
				AUC (95% CI)	<i>p</i>	H–L	<i>p</i>
POSSUM	10	1	10	0.76 (0.62–0.89)	0.005	2.96	0.94
P-POSSUM	10	2	5	0.85 (0.76–0.94)	0.001	8.29	0.41
CR-POSSUM	10	0	-	0.79 (0.67–0.90)	0.002	5.72	0.68
ACPGBI CRC	10	2	5	0.78 (0.59–0.97)	0.003	5.28	0.73
SORT	10	1	10	0.81 (0.68–0.94)	0.001	2.82	0.83

O observed, *E* expected, *AUC* area under curve, *95% CI* 95% confidence intervals, *H–L* hosmer–lemeshow

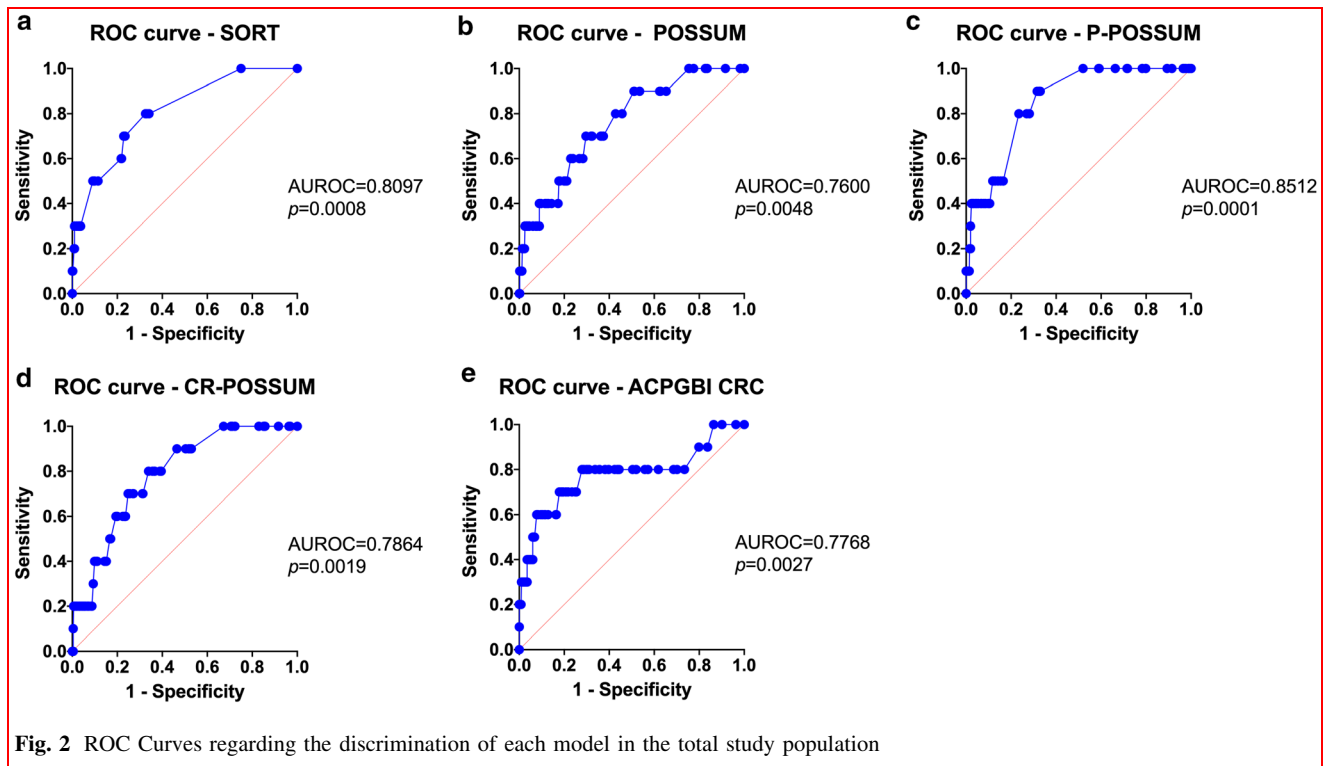


Fig. 2 ROC Curves regarding the discrimination of each model in the total study population

colon cancer (Fig. 3). In all other subgroups SORT was associated with fair discrimination attributes. Moreover, SORT demonstrated a high level of calibration in all subgroups, with the lowest value observed in patients undergoing open surgery. In fact, ACPGBI CRC model provided the best performing discrimination [AUC: 0.96 (0.91–1.00); $p < 0.001$] and calibration (H–L: 2.82; $p = 0.396$) for patients undergoing open surgery. In addition, it showed the highest discrimination level in patients undergoing acute surgery. In contrast, the accuracy of ACPGBI CRC to predict perioperative mortality was poor to fair in all other subgroups. P-POSSUM provided the best performing discrimination [AUC: 0.91 (95% CI: 0.80–1.00); $p = 0.002$] and calibration (H–L: 1.59; $p = 0.979$) in patients undergoing laparoscopic surgery. POSSUM and CR-POSSUM were associated with the worst discrimination-calibration balance in all subgroups. All models underpredicted perioperative mortality in all subgroups.

Discussion

The present study is the first to evaluate the validity of SORT model in (1) CRC, (2) non-UK patients undergoing surgery, (3) compare it with other risk models and (4) performing sensitivity subgroup analysis. The outcomes resulted from the current study might have a direct impact

in clinical practice, suggesting a possible role of SORT in the perioperative pathway and during the shared decision-making process of CRC patients.

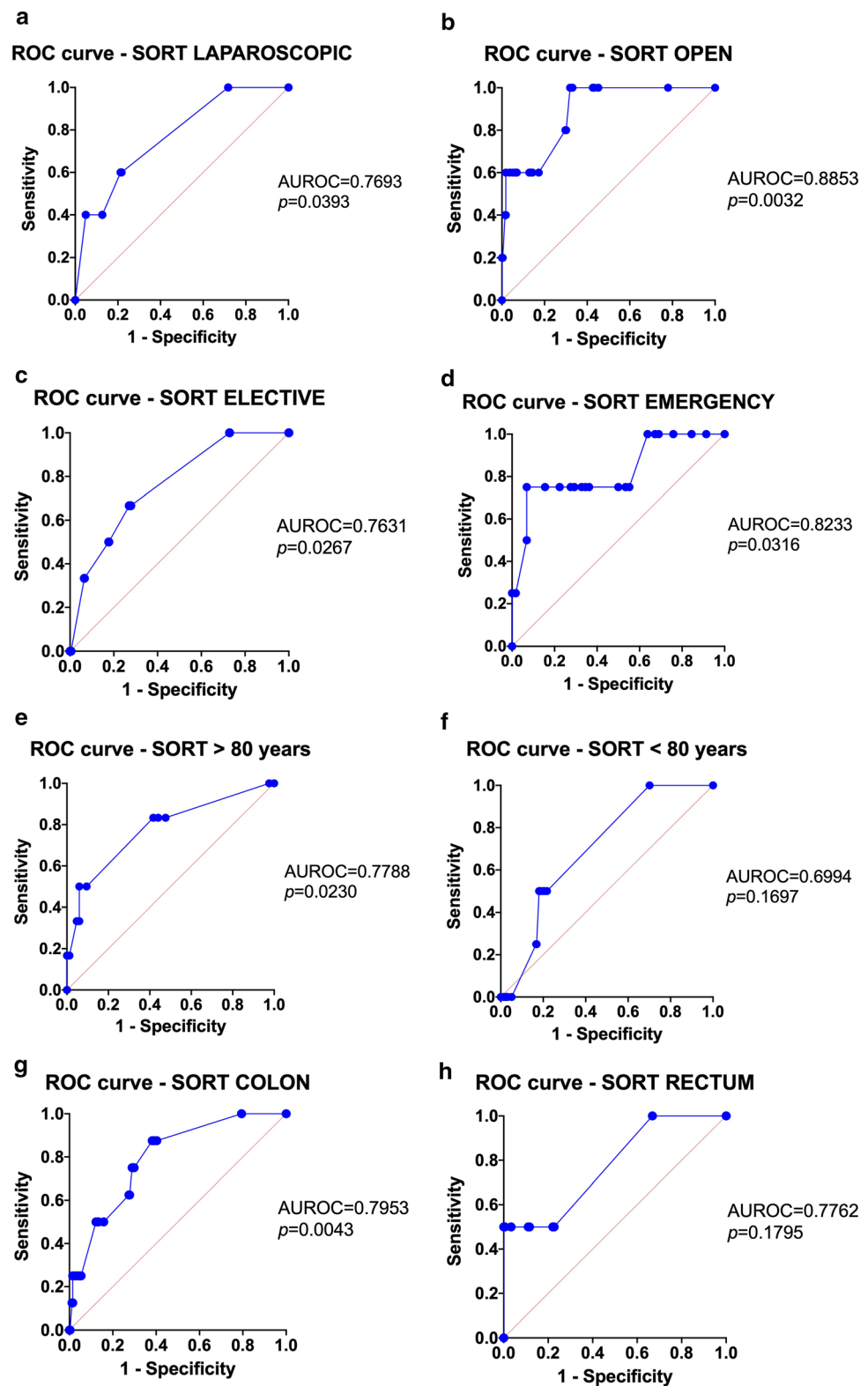
The SORT scoring system is a useful tool proposed by Protopapa et al. [3], to predict 30-day postoperative mortality. The study outcomes showed that six preoperatively available factors efficiently and effectively predicted postoperative mortality with a higher accuracy compared to other traditional risk assessment tools, such as ASA-PS [3]. Other risk stratification models that have been used in clinical practice and were included for comparison in the current study are POSSUM, P-POSSUM, CR-POSSUM, and ACPGBI CRC. Since, both patients and clinicians have implemented these tools in the counseling process, it was crucial to compare them with SORT. In addition, according to a recent study [13], POSSUM, P-POSSUM, CR-POSSUM and ACPGBI CRC were associated with poor accuracy in the setting of CRC surgery. The same study concluded that new models are required based on prospectively collected data [13]. Our outcomes provided an answer to this call. In fact, SORT demonstrated the best performing discrimination and calibration compared with all other risk stratification models assessed in the present study. All models underpredicted mortality. In this context, the study's outcomes have significant implications during counseling CRC patients regarding the perioperative mortality risk in order to decide the treatment strategy.

Table 3 Discrimination and calibration of the studied scores for predicting mortality in colorectal surgical subpopulations

Scoring systems	<i>O</i>	<i>E</i>	<i>O:E</i>	Discrimination		Calibration	
				AUC (95% CI)	<i>p</i>	H–L	<i>p</i>
<i>Open (n = 277)</i>							
POSSUM	5	1	5	0.84 (0.69–1.00)	0.008	5.38	0.717
P-POSSUM	5	2	2.5	0.89 (0.75–1.00)	0.003	5.07	0.750
CR-POSSUM	5	0	–	0.81 (0.64–0.99)	0.016	6.26	0.618
ACPGBI CRC	5	2	2.5	0.96 (0.91–1.00)	< 0.001	2.82	0.945
SORT	5	1	5	0.89 (0.77–1.00)	0.003	7.32	0.396
<i>Laparoscopy (n = 249)</i>							
POSSUM	5	0	–	0.70 (0.49–0.92)	0.117	4.40	0.820
P-POSSUM	5	0	–	0.91 (0.80–1.00)	0.002	1.59	0.979
CR-POSSUM	5	0	–	0.81 (0.66–0.95)	0.018	6.10	0.636
ACPGBI CRC	5	0	–	0.65 (0.36–0.94)	0.254	8.14	0.420
SORT	5	0	–	0.77 (0.57–0.97)	0.039	2.44	0.785
<i>Elective (n = 464)</i>							
POSSUM	6	0	–	0.69 (0.52–0.86)	0.112	3.09	0.93
P-POSSUM	6	0	–	0.83 (0.74–0.92)	0.006	14.92	0.061
CR-POSSUM	6	0	–	0.74 (0.60–0.88)	0.042	3.23	0.919
ACPGBI CRC	6	0	–	0.68 (0.41–0.94)	0.138	5.39	0.715
SORT	6	0	–	0.76 (0.59–0.94)	0.027	1.5	0.913
<i>Acute (n = 62)</i>							
POSSUM	4	1	4	0.78 (0.53–1.00)	0.059	10.49	0.233
P-POSSUM	4	1	4	0.89 (0.80–0.98)	0.010	8.19	0.415
CR-POSSUM	4	0	–	0.74 (0.51–0.97)	0.112	5.95	0.653
ACPGBI CRC	4	2	2	0.92 (0.80–1.00)	0.001	6.61	0.579
SORT	4	1	4	0.82 (0.58–1.00)	0.032	4.94	0.764
<i>Colon (n = 343)</i>							
POSSUM	8	0	–	0.75 (0.62–0.88)	0.016	4.19	0.840
P-POSSUM	8	0	–	0.83 (0.74–0.92)	0.001	9.81	0.279
CR-POSSUM	8	0	–	0.74 (0.60–0.89)	0.020	9.35	0.314
ACPGBI CRC	8	1	8	0.78 (0.58–0.98)	0.007	3.73	0.881
SORT	8	0	–	0.80 (0.66–0.93)	0.004	3.83	0.799
<i>Rectum (n = 183)</i>							
POSSUM	2	1	2	0.68 (0.24–1.00)	0.376	5.03	0.754
P-POSSUM	2	1	2	0.84 (0.61–1.00)	0.096	3.50	0.899
CR-POSSUM	2	1	2	0.89 (0.73–1.00)	0.060	9.45	0.306
ACPGBI CRC	2	1	2	0.68 (0.23–1.00)	0.383	6.24	0.621
SORT	2	1	2	0.78 (0.44–1.00)	0.180	1.42	0.922
<i>≥ 80 (n = 90)</i>							
POSSUM	6	1	6	0.78 (0.63–0.93)	0.023	13.43	0.098
P-POSSUM	6	2	3	0.85 (0.72–0.98)	0.004	6.43	0.490
CR-POSSUM	6	0	–	0.61 (0.35–0.87)	0.365	7.18	0.517
ACPGBI CRC	6	2	3	0.76 (0.53–0.99)	0.033	8.83	0.357
SORT	6	1	6	0.78 (0.56–0.99)	0.023	0.37	0.996
<i>< 80 (n = 436)</i>							
POSSUM	4	0	–	0.63 (0.40–0.85)	0.386	4.71	0.790
P-POSSUM	4	0	–	0.79 (0.66–0.93)	0.045	5.94	0.654
CR-POSSUM	4	0	–	0.71 (0.540.89)	0.140	4.38	0.820
ACPGBI CRC	4	0	–	0.60 (0.23–0.96)	0.495	7.16	0.519
SORT	4	0	–	0.70 (0.51–0.89)	0.170	4.59	0.468

O observed, *E* expected, *AUC* area under curve, *95% CI* 95% confidence intervals, *H–L* Hosmer–Lemeshow

Fig. 3 ROC Curves regarding the discrimination of SORT in each study subgroup



The feasibility of SORT was also proved in the sensitivity subgroup analysis. SORT was associated with fair to excellent discrimination and improved calibration.

Nonetheless, we should stress our comparative outcomes regarding two subgroups with direct clinical impact: ACPGBI CRC demonstrated the best performing

discrimination and calibration values for patients undergoing open surgery, while, P-POSSUM demonstrated the best performing discrimination and calibration values for patients having laparoscopic surgery. This observation might be of great importance in daily practice, during the counseling and shared-decision making process regarding the optimal surgical approach (laparoscopic or open) for a given patient. In addition, SORT was associated with a moderate-high level of discrimination in patients aged ≥ 80 years. In fact, this observation is clinically important, since age and preoperative frailty are associated with postoperative morbidity in patients undergoing surgery for colorectal cancer [14, 15]. Furthermore, the findings of the present study regarding the value of clinical variables employed by SORT, are in accordance with evidence provided by administrative datasets [16]. Besides, according to our outcomes SORT presents higher accuracy compared with other preoperative (Barwon Health 2009–BH 2009) [17], along with intraoperative risk stratification models (Surgical Apgar Score) [18], while remaining friendly-to-use, since it implements only six preoperative variables.

ACPGBI CRC has been previously validated in CRC surgical patients [19] and demonstrated poor predicting power. Our findings are in accordance with this study [19], with the exception of open surgical approach subgroup. Even though, P-POSSUM has been extensively validated [2], the SORT has a number of advantages over it. First of all, SORT incorporates only six preoperative variables, compared with eighteen perioperative variables of P-POSSUM, thus being significantly easier to be implemented in the daily clinical practice. Secondly, P-POSSUM includes intra- and postoperative variables that are not available during the preoperative assessment. Finally, P-POSSUM contains subjective variables, thus increasing the interobserver variability and heterogeneity and posing a certain bias in its predicting accuracy.

A certain limitation of the present study has to do with the design, as it is a single-institution, retrospective study. However, the data are prospectively collected, the patients are consecutive, the surgical team is the same and the surgeon's bias regarding patient/approach selection was minimized, as this was mostly depending on patients' preference and/or logistics. Another potential bias is the belief of many patients in Greece that laparoscopic colectomy is still performed at an "experimental" level. To reduce this bias there has been an extensive counseling with each patient in order to choose the desired surgical strategy through a shared decision-making process.

The current outcomes demonstrate that SORT is an easy, feasible and efficient risk stratification tool that should be implemented in the preoperative counseling and shared decision-making process of CRC patients.

Conclusion

In this study we validated the SORT risk stratification model in Greek adult patients undergoing surgery for colorectal cancer. SORT demonstrated the best performing discrimination and calibration compared with POSSUM, P-POSSUM, CR-POSSUM, and ACPGBI CRC. The value of SORT was further confirmed by sensitivity subgroup analysis. SORT is a feasible and efficient risk stratification tool that could be implemented in the perioperative pathway of CRC patients.

Acknowledgements The present study was originally written as part of a UCL Master's degree within the UCL Centre for Perioperative Medicine.

Author contributions DEM contributed to protocol development, acquisition of data, analysis and interpretation of data, drafting the article or revising it critically for important intellectual content, and final approval of the version to be published. DW contributed to protocol development, acquisition of data, interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published. IB contributed to conception and design, acquisition of data, interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published. MF contributed to acquisition of data, interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published. IM contributed to acquisition of data, interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published. GC contributed to acquisition of data, drafting the article, and final approval of the version to be published. GT contributed to protocol development, acquisition of data, analysis and interpretation of data, drafting the article or revising it critically for important intellectual content, and final approval of the version to be published.

Compliance with ethical standards

Conflicts of interest All authors declare no conflicts of interest regarding the present study. This research did not receive grants from any funding agency in the public, commercial or not-for-profit sectors.

Ethical approval Ethical approval was obtained by the Scientific Committee of the University Hospital of Larissa (Protocol number: 33606/16–07-19).

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. *CA Cancer J Clin* 65:87–108. <https://doi.org/10.3322/caac.21262>
2. Moonesinghe SR, Mythen MG, Das P, Rowan KM, Grocott MP (2013) Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: qualitative systematic review. *Anesthesiology* 119(4):959–981. <https://doi.org/10.1097/ALN.0b013e3182a4e94d>

3. Protopapa KL, Simpson JC, Smith NC, Moonesinghe SR (2014) Development and validation of the surgical outcome risk tool (SORT). *Br J Surg* 101(13):1774–1783. <https://doi.org/10.1002/bjs.9638>
4. Marufu TC, White SM, Griffiths R, Moonesinghe SR, Moppett IK (2016) Prediction of 30-day mortality after hip fracture surgery by the nottingham hip fracture score and the surgical outcome risk tool. *Anaesthesia* 71(5):515–521. <https://doi.org/10.1111/anae.13418>
5. Wong GTC, Ang WC, Wong TCL, Choi SW (2017) Surgical outcome risk tool (SORT) validation in hepatectomy. *Anaesthesia* 72(10):1287–1289
6. Copeland GP, Jones D, Walters M (1991) POSSUM: a scoring system for surgical audit. *Br J Surg* 78(3):355–60.22
7. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ (1998) POSSUM and portsmouth POSSUM for predicting mortality. Physiological and operative severity score for the enUmeration of mortality and morbidity. *Br J Surg* 85(9):1217–1220
8. Tekkis PP, Prytherch DR, Kocher HM, Senapati A, Poloniecki JD, Stamatakis JD, Windsor AC (2004) Development of a dedicated risk-adjustment scoring system for colorectal surgery (colorectal POSSUM). *Br J Surg* 91(9):1174–1182
9. Ferjani AM, Griffin D, Stallard N, Wong LS (2007) A newly devised scoring system for prediction of mortality in patients with colorectal cancer: a prospective study. *Lancet Oncol* 8(4):317–322
10. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837–845
11. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S (1997) A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 16:965–980
12. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP (2007) STROBE initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 147(8):573–577 (**Erratum in: *Ann Intern Med.* 2008 Jan 15;148(2):168. PMID: 17938396**)
13. Bagnall NM, Pring ET, Malietzis G et al (2018) Perioperative risk prediction in the era of enhanced recovery: a comparison of POSSUM, ACPGIBI, and E-PASS scoring systems in major surgical procedures of the colorectal surgeon. *Int J Colorectal Dis* 33:1627–1634. <https://doi.org/10.1007/s00384-018-3141-4>
14. Richards SJG, Cherry TJ, Frizelle FA, Eglinton TW (2020) Preoperative frailty is predictive of adverse post-operative outcomes in colorectal cancer patients. *ANZ J Surg*. <https://doi.org/10.1111/ans.16319>
15. Ragg JL, Watters DA, Guest GD (2009) Preoperative risk stratification for mortality and major morbidity in major colorectal surgery. *Dis Colon Rectum* 52(7):1296–1303. <https://doi.org/10.1007/DCR.0b013e3181a0e639>
16. García-Torrecillas JM, Olvera-Porcel MC, Ferrer-Márquez M et al (2020) Predictive model of the risk of in-hospital mortality in colorectal cancer surgery, based on the minimum basic data set. *Int J Environ Res Public Health* 17(12):4216. <https://doi.org/10.3390/ijerph17124216>
17. Regenbogen SE, Bordeianou L, Hutter MM, Gawande AA (2010) The intraoperative surgical appar score predicts postdischarge complications after colon and rectal resection. *Surgery* 148(3):559–566. <https://doi.org/10.1016/j.surg.2010.01.015> (**Epub 2010 Mar 12**)
18. Kong CH, Guest GD, Stupart DA, Faragher IG, Chan ST, Watters DA (2013) Recalibration and validation of a preoperative risk prediction model for mortality in major colorectal surgery. *Dis Colon Rectum* 56(7):844–849. <https://doi.org/10.1097/DCR.0b013e31828343f2>
19. Degett TH, Roikjær O, Iversen LH, Gögenur I (2018) A model predicting operative mortality in the UK has only limited value in Denmark. *Int J Colorectal Dis* 33:141–147. <https://doi.org/10.1007/s00384-017-2937-y>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.