Accuracy of Risk Estimation for Surgeons *Versus* Risk Calculators in Emergency General Surgery

Lauren V. Huckaby, MD, MS,

Esmaeel Reza Dadashzadeh, MD, MS,

Shimena Li, MD,

Insiyah Campwala, BS,

Lucine Gabriel, BS,

Jason Sperry, MD, MPH,

Robert M. Handzel, MD, MS,

Raquel Forsythe, MD,

Joshua Brown, MD, MSc[*]

Department of Surgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

Abstract

Introduction: Surgical risk calculators have expanded in both number and sophistication of their predictive approach. These calculators are gaining popularity as validated tools to help surgeons estimate mortality and complications following emergency general surgery (EGS). However, the accuracy of risk estimates generated by these calculators compared to risk estimation by practicing surgeons has not been explored.

Methods: Acute care surgeons at a quaternary care center prospectively estimated 30-d mortality and complications for adult EGS patients (2019-2021). Surgeon predictions were compared to Predictive OpTimal Trees in Emergency Surgery Risk (POTTER) and NSQIP estimates. Observed-to-expected (O:E) ratios of median aggregate estimates were calculated. C-statistics for surgeon and calculator estimations were utilized to quantify predictive accuracy.

Results: Among 150 patients (median 61 y, 45% male), 30-d mortality was 15% ($n = 23$). Observed rates of prolonged mechanical ventilation and acute renal failures were 30% and 10%, respectively. Overall, surgeon predictions were similar to risk calculator estimates for mortality (c-statistics 0.843 [surgeon] *versus* 0.848 [POTTER] and 0.815 [NSQIP]) and need

---

[*]*Corresponding author.* Division of General Surgery and Trauma, Department of Surgery, UPMC Presbyterian, 200 Lothrop Street, Pittsburgh, PA 15213. Tel.: +1 412 802 8269; fax: +1 412 647 1448. brownjb@upmc.edu (J. Brown).

for prolonged ventilation (c-statistics 0.801 *versus* 0.722 and 0.689, respectively). Surgeons tended to overestimate complication risks. Surgeon experience was not significantly associated with mortality prediction in an adjusted model.

Conclusions: Acute care surgeons at a quaternary care center predicted postoperative mortality and complications with similar discrimination when compared to surgical risk calculators. Surgeon expertise should be utilized in conjunction with risk calculators when counseling EGS patients regarding anticipated postoperative outcomes. Surgeons should be cognizant of patterns in overestimation or underestimation of complications.

Keywords

Emergency general surgery; Postoperative complications; Risk calculators; Surgical risk stratification

Introduction

Informed consent and discussion of the anticipated postoperative course are key components of patient management for emergency general surgery (EGS) cases. However, given the medical and surgical complexity of some patients undergoing operative intervention, it may be difficult to accurately predict postoperative outcomes. This is particularly true at quaternary care centers, in which patients are often transferred for EGS assessment due to the complexity of their presentation or comorbidities. Thus, evidence-based surgical risk calculators have been developed to provide objective estimates of surgical morbidity and mortality.[1-4] The American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP) surgical risk calculator is one such tool that was developed utilizing a large surgical database.[3,5,6] More recently, Bertsimas *et al.* reported on results utilizing the Predictive OpTimal Trees in Emergency Surgery Risk (POTTER) calculator, which was derived from EGS patients.[2] Multiple other surgical calculators have also been described.[3,7-10] These calculators span various surgical subspecialties and differ in their derivation, validation, and accuracy. Nevertheless, these calculators are intended to be utilized by experienced surgeons, and the results of these predictions should be communicated to patients as one component of a broad discussion of surgical risk. The experience of the operating surgeon may also influence this discussion, and thus far, there have been no studies comparing the accuracy of surgeon risk estimation with these new tools. We hypothesized that surgeons at a large volume academic medical center would be able to predict the risk of mortality and complications with comparable accuracy when compared to the POTTER and ACS-NSQIP calculators.

Methods

Study design

Full details of the study design are provided in the Supplemental Methods. Briefly, attending surgeons ($n = 17$) in the Division of Trauma Surgery at a large quaternary care academic medical center were asked to prospectively quantify 30-d risks of mortality, need for mechanical ventilation >48 h, acute renal failure (ARF, defined as the need for renal replacement therapy), and intra-abdominal infection (defined as radiologic findings of an

intra-abdominal abscess) for adult (18 y) EGS patients. Risk predictions were recorded prior to the commencement of the operation by the operating surgeon. A retrospective chart review was then utilized to obtain patient characteristics and 30-d outcomes. This study was approved by the Institutional Review Board at the University of Pittsburgh (#18100079).

Calculation of risk scores

The POTTER risk score was calculated using the POTTER software application (version 1.1, Alexandria Health).[2] NSQIP scores were calculated using the online calculator (American College of Surgeons, Chicago, IL).[3] A detailed description of the score calculation is provided in Supplemental Methods.

Statistical analysis

Continuous data are presented as mean (standard deviation) or median (interquartile range [IQR]) as appropriate. Categorical data are presented as numbers (percentages). Given the non-Gaussian distribution of risk predictions, discussion of the results refers to only the medians of risk predictions; however, the means are provided in the tables for comparison. Observed-to-expected (O:E) ratios were calculated by dividing the observed rates of mortality or complications by either the median or mean of the risk estimations, as indicated in the table. The predicted risk of mortality and complications for surgeons, POTTER and NSQIP, was compared to the observed rates using both the C-statistic and Brier score. The C-statistic is a measure of discrimination, while the Brier score is a measure of the accuracy of prediction. To account for missing laboratory values in some patients, we performed a separate analysis using normal laboratory values in order to calculate a risk score (Supplemental Methods). Receiver operating curves were constructed for the estimation of mortality. A multivariable linear regression model, including patient characteristics and surgeon experience (i.e., <5 y of post-training experience compared to those with 5 y of experience), was constructed to model independent predictors of a higher estimated mortality rate. Surgeon experience as a binary variable was included as a forced variable. Stepwise regression was utilized to construct a final model, including all variables with $P < 0.05$. All hypothesis testing was two-sided. Statistical analyses were performed using the Stata 16 software package (StataCorp, 2017, Stata Statistical Software: Release 16, College Station, TX).

Results

Study population

A total of 150 patients (median age 61 y, 45% male) were identified, with 122 (81%) undergoing intra-abdominal operations (Table 1, Supplemental Table 1). The median body mass index was 29.3 kg/m^2, and 42 patients (28%) were current smokers. Additionally, a majority of patients ($n = 81$, 54%) carried a diagnosis of hypertension, and 39 patients (26%) had diabetes. The median creatinine level was 1.03 mg/dL (IQR 0.80-1.79 mg/dL), and the median white blood cell count was 11.7×10^9/L (IQR $8.3\text{-}18.1 \times 10^9$/L). The median American Society of Anesthesiologists (ASA) Classification was 3 (IQR 3-4).

Comparison of observed and expected outcomes

Observed 30-d mortality was 15% ($n = 23$) (Table 2). The median estimated mortality from surgeon prediction was 8.5% (IQR 1%-25%) compared to 4.8% (IQR 1%-21%) for POTTER and 7% (IQR 2%-16%) for NSQIP. Using these medians, both surgeons and the risk calculators underestimated the mortality risk (O:E ratio 1.80 for surgeons and 3.23 *versus* 2.22 for POTTER and NSQIP, respectively). C-statistics for mortality were 0.843 for surgeons compared to 0.848 and 0.815 for POTTER and NSQIP, respectively. The Brier scores were low for all three estimates, indicating good accuracy of prediction. Receiver operating curves are shown for mortality (Fig.).

Overall, 14 of 140 (10%) susceptible patients demonstrated acute renal failure. Median estimated rates of acute renal failure were 10% for surgeons (IQR 2%-20%) compared to 1% for POTTER (IQR 0%-4%) and 2% (IQR 1%-4%) for NSQIP. The surgeon's risk prediction was identical to the observed rate of ARF while both calculators underestimated the risk of ARF (O:E ratios of 10.42 and 4.44 for POTTER and NSQIP, respectively). C-statistics were 0.801, 0.722, and 0.689 for surgeons, POTTER, and NSQIP, respectively.

The observed rate of prolonged mechanical ventilation was 30% ($n = 45$). The median predicted rate was 10% for surgeons (IQR 2%-60%) as compared to 10% for POTTER (IQR 2%-21%). C-statistics were 0.874 and 0.876, respectively. The observed rate of intra-abdominal infection was 9% ($n = 11$ of 122 patients). Median predicted rates were 15% (IQR 5%-25%) for surgeons and 3% (IQR 2%-4%) for POTTER. C-statistics were 0.514 and 0.625, respectively.

Observed and predicted outcomes, utilizing normal lab values for those patients in whom missing lab values precluded score calculation, are shown (Supplemental Table 2).

Analysis of surgeon experience and predicted risk of mortality

We performed a subanalysis stratified by surgeon ($n = 17$) experience (i.e., <5 y or 5 y of post-training practice). The majority of cases ($n = 95$, 63%) in this series were performed by surgeons with 5 y of experience. Both groups of surgeons underestimated mortality risk. A multivariable linear regression model of predictors influencing an elevated predicted risk of mortality is shown. Surgeon experience was not significantly associated with the predicted risk of mortality. Patient age, the presence of hypertension, ascites, sepsis, and lab values for the white blood cell count and albumin level were all associated with mortality prediction (Table 3).

Discussion

Progress in statistical and computing methodology has allowed for the application of artificial intelligence principles to estimate surgical risk. This has resulted in the development of various surgical risk calculators, which are variably applied to general and subspecialty surgical populations. Nevertheless, there is a continued need to improve upon estimation and communication of surgical risks to patients and families, particularly as the complexity of operative cases continues to evolve. In this study, we prospectively queried surgeons for their predictions of mortality and complications for emergency general surgery

patients. We found that, overall, surgeon predictions were similar to those estimations generated by risk calculators, although there were distinct patterns in underestimation and overestimation of risks. In an adjusted model, surgeon experience was not significantly associated with estimates of mortality risk. This work has important implications in the understanding of the utility of surgical risk calculators alongside surgeon expertise and calls for further work into areas of potential improvement for surgical risk estimation.

We found comparable accuracy between aggregate surgeon predictions when compared to risk calculator estimations for EGS patients. In the original description of the POTTER calculator, the authors reported a C-statistic for 30-d mortality of 0.898 for POTTER as compared to 0.874 for the ASA estimation, 0.891 for the Emergency Surgery Score, and 0.898 for the NSQIP calculator.[2] In comparison, our C-statistics for 30-d mortality were 0.843 for surgeons and 0.848 for POTTER. Surgeons in our study were also able to predict the need for prolonged mechanical ventilation and acute renal failure with high accuracy. Prior work has described the increased morbidity and mortality among EGS patients.[11-15] Havens *et al.*, for example, found that emergency surgery remained an independent risk factor for death and complications when compared to nonemergency surgery, which persisted following adjustment for baseline characteristics and the type of operation.[11] Accurate prediction of mortality is perhaps the most critical outcome for surgeons, and thus, one of the primary aims of this study was to improve understanding of variations in mortality estimation. We were reassured that surgeon discrimination was similar to calculator estimates, particularly given that the vast majority of surgeons at our institution do not routinely utilize risk calculators. We do acknowledge that risk prediction among surgeons may vary based on patient-specific factors, including the type of case, and thus, a further inquiry would be necessary to determine differential risk prediction among surgical subpopulations.

Although we focused only on the POTTER and NSQIP calculators, a multitude of risk calculators for various surgical disciplines exist.[10,16] There have also been attempts to extrapolate risk estimations in different patient populations.[17-19] The availability of multiple risk calculators raises the question of which estimation approach is optimal. The POTTER calculator was recently developed to address risk prediction specifically in the EGS population and has since been validated in the elderly EGS population.[2,8] This tool represents a new era of risk prediction approaches utilizing machine learning techniques.[20] We found that the POTTER calculator displayed higher discrimination when compared to NSQIP for our patient population. This highlights the need for risk prediction models, which account for factors such as the urgency of the operative intervention. Integration of risk calculators into the electronic health records may further facilitate estimation and promote the use of these tools.

In this study, we specifically examined the high-volume quaternary care setting to capture predictions among the most critically ill patient groups. Prior work from our institution has demonstrated in-hospital mortality rates of up to 37% for EGS cases originating from medical intensive care unit consultations.[21] Thus, it was not unexpected that we found a mortality rate of 15% for the population described here. In fact, of our 150 cases, there were only 6 (4%) laparoscopic cholecystectomies and 7 (5%) laparoscopic

appendectomies, which is reflective of our typical volume of complex cases, many of which can be classified as surgical rescue.[22] A significant proportion of our patients are transferred from other facilities, and thus, surgeon predictions may account for factors, which may not be captured by surgical risk calculators, including delay in diagnosis or delay in transfer. As a large transplant center, many of these EGS operations are also performed on transplant patients, the nuances of which may not be fully captured by the calculator variable as 'immunocompromised'.[23] The NSQIP calculator providers a modifier for 'Surgeon Adjustment of Risks', allowing for classification of a patient risk being 'somewhat higher' or 'significantly higher'.[24] Such an adjustment is not currently available within the POTTER calculator. Although a preliminary step, this correction highlights the importance of surgeon interpretation of patient risk in combination with calculator output. Further work will be critical in attempting to objectify some of these unmeasured risks and understanding the best approach to incorporate surgeon input into these models.

Collectively, our results bring to attention the role of risk estimation in other practice settings. Prior work has demonstrated that lower hospital volume is associated with worse outcomes for EGS cases and that EGS outcomes vary widely by hospital setting.[14,25] Thus, risk estimations may also assist in decision-making to transfer a patient to a higher volume center for continued care. Future work from our group will seek to explore the accuracy of risk predictions among surgeons performing EGS cases at lower volume hospitals. We also intend to explore risk prediction among surgeons who have and have not completed advanced training (i.e., fellowship). It is plausible that risk calculators may serve a unique role in assisting in the optimal triage of patients in resource-limited settings in the future.

The most important application of this work is an improved understanding of the approach to counseling patients, especially critically ill patients and their families, with regard to the relative risks of surgery.[26] We strongly believe that risk calculators will evolve and play an increasingly important role in the discussion of surgical risks, particularly as advances in risk prediction techniques improve the accuracy of these estimates.[27,28] Nevertheless, these estimates remain only one component of the surgical discussion. These calculators were not developed with the intent of having patients calculate their own risks, and thus, the explanation and interpretation provided by a trained surgeon are essential to contextualize these numbers.

This study has several limitations. First and foremost, this work was conducted at a large academic medical center, and our findings should not be applied to lower volume centers until validity in these settings has been evaluated. We acknowledge that we were unable to capture all emergency general surgery cases that occurred during this time period; however, we feel that the breadth of cases that were captured is reflective of the typical case variety and complexity at our institution. We also do not account for patients who were not offered an operation or who declined surgery. Finally, we did not specifically explore inter-rater reliability, and there may be unmeasured factors that influenced surgeon predictions, which were not accounted for. Prior work performed at our institution has reported on institutional resources and professional norms as contributors to decision-making for EGS cases[29]; the role of these factors in predicting EGS outcomes merits further investigation.

Conclusions

We report on the results of a prospective examination of surgeon prediction of postoperative risks for patients undergoing EGS cases. We found that surgeons at a high volume tertiary care center were able to predict complication rates with high accuracy compared to observed outcomes and furthermore that these surgeon predictions were comparable to those provided by risk calculators. Further work is necessary to understand surgeon prediction in other settings (i.e., community-based practices, other patient subpopulations), particularly as risk calculators continue to evolve. We suggest that these risk calculators should continue to be utilized as a component of surgical risk discussions. Surgeon expertise and understanding of patient and family goals and wishes, however, should remain as key factors in the shared decision-making for surgical patient management.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Hyer JM, White S, Cloyd J, et al. Can we improve prediction of adverse surgical outcomes? Development of a surgical complexity score using a novel machine learning technique. J Am Coll Surg. 2019;230:43–52.e1. [PubMed: 31672674]

2. Bertsimas D, Dunn J, Velmahos GC, Kaafarani HMA. Surgical risk is not linear: derivation and validation of a novel, user-friendly, and machine-learning-based predictive optimal trees in emergency surgery risk (POTTER) calculator. Ann Surg. 2018;268:574–583. [PubMed: 30124479]

3. Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. J Am Coll Surg. 2013;217:833–842.e3. [PubMed: 24055383]

4. El Hechi M, Maurer LR, Levine J, et al. Validation of the artificial intelligence-based predictive optimal trees in emergency surgery risk (POTTER) calculator in emergency general surgery and emergency laparotomy patients. J Am Coll Surg. 2021;232:912–919. [PubMed: 33705983]

5. Cohen ME, Liu Y, Ko CY, Hall BL. An examination of American College of Surgeons NSQIP surgical risk calculator accuracy. J Am Coll Surg. 2017;224:787–795.e1. [PubMed: 28389191]

6. Liu Y, Cohen ME, Hall BL, Ko CY, Bilimoria KY. Evaluation and enhancement of calibration in the American College of Surgeons NSQIP surgical risk calculator. J Am Coll Surg. 2016;223:231–239. [PubMed: 27212006]

7. Shahian DM, Jacobs JP, Badhwar V, et al. The society of thoracic surgeons 2018 adult cardiac surgery risk models: Part 1—background, design considerations, and model development. Ann Thorac Surg. 2018;105:1411–1418. [PubMed: 29577925]

8. Maurer LR, Chetlur P, Zhuo D, et al. Validation of the AI-based predictive optimal trees in emergency surgery risk (POTTER) calculator in patients 65 years and older [e-pub ahead of print]. Ann Surg. 2020. 10.1097/sla.0000000000004714.

9. Hall DE, Arya S, Schmid KK, et al. Development and initial validation of the risk analysis index for measuring frailty in surgical populations. JAMA Surg. 2017;152:175–182. [PubMed: 27893030]

10. Nandan AR, Bohnen JD, Sangji NF, et al. The Emergency Surgery Score (ESS) accurately predicts the occurrence of postoperative complications in emergency surgery patients. J Trauma Acute Care Surg. 2017;83:84–89. [PubMed: 28422908]

11. Havens JM, Peetz AB, Do WS, et al. The excess morbidity and mortality of emergency general surgery. J Trauma Acute Care Surg. 2015;78:306–311. [PubMed: 25757115]

12. Akinbami F, Askari R, Steinberg J, Panizales M, Rogers SO. Factors affecting morbidity in emergency general surgery. Am J Surg. 2011;201:456–462. [PubMed: 21421099]

13. Huckaby LV, Dadashzadeh ER, Handzel R, Kacin A, Rosengart MR, van der Windt DJ. Improved understanding of acute incisional hernia incarceration: implications for addressing the excess mortality of emergent repair. J Am Coll Surg. 2020;231:536–546. [PubMed: 32822886]

14. Ogola GO, Crandall ML, Shafi S. Variations in outcomes of emergency general surgery patients across hospitals: a call to establish emergency general surgery quality improvement program. J Trauma Acute Care Surg. 2018;84:280–286. [PubMed: 29194319]

15. Gale SC, Shafi S, Dombrovskiy VY, Arumugam D, Crystal JS. The public health burden of emergency general surgery in the United States: a 10-year analysis of the Nationwide Inpatient Sample - 2001 to 2010. J Trauma Acute Care Surg. 2014;77:202–208. [PubMed: 25058242]

16. Stonelake S, Thomson P, Suggett N. Identification of the high risk emergency surgical patient: which risk prediction model should be used? Ann Med Surg. 2015;4:240–247.

17. Veeravagu A, Li A, Swinney C, et al. Predicting complication risk in spine surgery: a prospective analysis of a novel risk assessment tool. J Neurosurg Spine. 2017;27:81–91. [PubMed: 28430052]

18. Rivard C, Nahum R, Slagle E, Duininck M, Isaksson Vogel R, Teoh D. Evaluation of the performance of the ACS NSQIP surgical risk calculator in gynecologic oncology patients undergoing laparotomy. Gynecol Oncol. 2016;141:281–286. [PubMed: 26899020]

19. Prasad KG, Nelson BG, Deig CR, Schneider AL, Moore MG. ACS NSQIP risk calculator: an accurate predictor of complications in major head and neck surgery? Otolaryngol - Head Neck Surg (United States). 2016;155:740–742.

20. Jalali A, Lonsdale H, Do N, et al. Deep learning for improved risk prediction in surgical outcomes. Sci Rep. 2020;10:1–13. [PubMed: 31913322]

21. Briggs A, Handzel RM, Kutcher ME, Peitzman AB, Forsythe RM. Predisposed to failure? The challenge of rescue in the medical intensive care unit. J Trauma Acute Care Surg. 2019;87:774–781. [PubMed: 31233441]

22. Kutcher MEM, Sperry Jason L, et al. Surgical rescue: the next pillar of acute care surgery. J Trauma Acute Care Surg. 2017;82:280–286. [PubMed: 27893639]

23. Bhatti UF, Shah AA, Williams AM, et al. Defining the burden of emergency general surgery in transplant patients: a nationwide examination. J Surg Res. 2020;245:315–320. [PubMed: 31421379]

24. ACS NSQIP surgical risk calculator. American College of Surgeons. Available at: https://riskcalculator.facs.org/RiskCalculator/. Accessed December 9, 2021.

25. Mehta A, Efron DT, Canner JK, et al. Effect of surgeon and hospital volume on emergency general surgery outcomes. J Am Coll Surg. 2017;225:666–675.e2. [PubMed: 28838870]

26. Grady C. Enduring and emerging challenges of informed consent. N Engl J Med. 2015;372:855–862. [PubMed: 25714163]

27. Campbell G, Watters DAK. Making decisions in emergency surgery. ANZ J Surg. 2013;83:429–433. [PubMed: 23656408]

28. Paruch JL, Ko CY, Bilimoria KY. An opportunity to improve informed consent and shared decision making: the role of the ACS NSQIP surgical risk calculator in oncology. Ann Surg Oncol. 2014;21:5–7. [PubMed: 24197763]

29. Kulkarni SS, Briggs A, Sacks OA, et al. Inner deliberations of surgeons treating critically-ill emergency general surgery patients: a qualitative analysis. Ann Surg. 2021;274:1081–1088. [PubMed: 31714316]
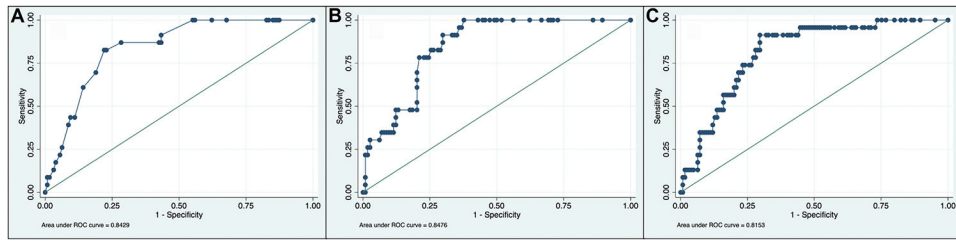
Fig. –.

Receiver operating curves for mortality estimates based on surgeon prediction (A), POTTER prediction (B), and NSQIP prediction (C).

Table 1 –

Baseline characteristics of patients undergoing emergency general surgery.

Characteristic	Emergency general surgery patients *n* = 150
Age (y) – median [IQR]	61 [50-70]
Male sex – no. (%)	67 (45)
Body mass index (kg/m^2) – median [IQR]	29.3 [23.7-36.1]
Current smoker – no. (%)	42 (28)
Cancer – no. (%)	17 (11)
Chronic obstructive pulmonary disease – no. (%)	22 (15)
Hypertension – no. (%)	81 (54)
Diabetes – no. (%)	39 (26)
Congestive heart failure – no. (%)	12 (8)
Dialysis – no. (%)	10 (7)
Immunocompromised – no. (%)	28 (19)
Sodium (mEq/L) – mean (SD)	136.67 (4.52)
Creatinine (mg/dL) – median [IQR]	1.03 [0.80-1.79]
Albumin (mg/dL) – mean (SD)	3.24 (0.81)
Total bilirubin (mg/dL) – median [IQR]	0.60 [0.50-1.00]
White blood cell count (×10^9/L)– median [IQR]	11.70 [8.30-18.10]
Platelet count (×10^9/L) – median [IQR]	246 [184-309]
International normalized ratio – median [IQR]	1.2 [1.1-1.5]
ASA classification – median [IQR]	3 [3-4]

Table 2 –

Comparison of observed and predicted outcomes at 30 d in patients undergoing emergency general surgery cases.

Characteristic	Observed no. (%)	Surgeon prediction	POTTER estimation	NSQIP estimation
Mortality	23 (15)			
Mean % (SD)		17 (23)	13 (15)	13 (17)
Median % [IQR]		9 [1-25]	5 [1-21]	7 [2-16]
O:E ratio (mean)		0.89	1.21	1.18
O:E ratio (median)		1.80	3.23	2.22
C-statistic (95% CI)		0.843 (0.771-0.915)	0.848 (0.780-0.914)	0.815 (0.734-0.897)
Brier score		0.114	0.112	0.115
Mechanical Ventilation >48 h	45 (30)			
Mean % (SD)		30 (36)	17 (19)	–
Median % [IQR]		10 [2-60]	10 [2-22]	–
O:E ratio (mean)		1.00	1.82	–
O:E ratio (median)		3.00	3.10	–
C-statistic (95% CI)		0.874 (0.807-0.942)	0.876 (0.815-0.937)	–
Brier score		0.126	0.175	–
Renal failure	14/140 (10)			
Mean % (SD)		19 (27)	3 (6)	4 (4)
Median % [IQR]		10 [2-20]	1 [0-4]	2 [1-4]
O:E ratio (mean)		0.52	3.22	2.81
O:E ratio (median)		1.00	10.42	4.44
C-statistic (95% CI)		0.801 (0.660-0.941)	0.722 (0.561-0.884)	0.689 (0.567-0.811)
Brier score		0.093	0.068	0.077
Intra-abdominal infection	11/122 (9)			
Mean % (SD)		21 (21)	4 (3)	–
Median % [IQR]		15 [5-25]	3 [2-4]	–
O:E ratio (mean)		0.43	2.52	–
O:E ratio (median)		0.60	2.91	–
C-statistic (95% CI)		0.514 (0.298-0.725)	0.625 (0.410-0.840)	–
Brier score		0.140	0.086	–

Table 3 –

Multivariable linear regression for prediction of elevated mortality by surgeons. Surgeon experience ($<$ 5 y or ≥ 5 y) was included as a forced variable.

Characteristic	Coefficient	95% confidence interval	P value
Surgeon experience ≥ 5 y	−3.61	−9.99 to 2.77	0.265
Age	0.28	0.07 to 0.50	0.011
Hypertension	−7.07	−13.64 to −0.50	0.035
Ascites	19.74	11.49 to 27.99	<0.001
Sepsis	16.82	6.79 to 26.85	0.001
White blood cell count	0.70	0.26 to 1.14	0.002
Albumin level	−6.81	−11.13 to −2.48	0.002

Variables included in the model: surgeon experience, age, sex, body mass index, cancer, dyspnea, current smoker, chronic obstructive pulmonary disease, hypertension, diabetes, congestive heart failure, dialysis, sepsis, bleeding disease, immunocompromise, ascites, albumin level, white blood cell count level.