Medicine

OPEN

Assessing the effectiveness of ACS surgical risk calculator versus P-POSSUM in predicting mortality and morbidity for major hepatobiliary surgery

An observational study

Ali Karabulut, MD[a], Veysel Umman, MD[b], Gunes Oral, MD[a], Ergin Erginoz, MD[a,*], Mehmet Sinan Carkman, MD[a]

Abstract

Risk assessment is difficult yet would provide valuable data for both the surgeons and the patients in major hepatobiliary surgeries. An ideal risk calculator should improve workflow through efficient, timely, and accurate risk stratification. The American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) surgical risk calculator (SRC) and Portsmouth Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (P-POSSUM) are surgical risk stratification tools used to assess postoperative morbidity. In this study, preoperative data from 300 patients undergoing major hepatobiliary surgeries performed at a single tertiary university hospital were retrospectively collected from electronic patient records and entered into the ACS-SRC and P-POSSUM systems, and the resulting risk scores were calculated and recorded accordingly. The ACS-NSQIP-M1 (C-statistics = 0.725) and M2 (C-statistics = 0.791) models showed better morbidity discrimination ability than P-POSSUM-M1 (C-statistics = 0.672) model. The P-POSSUM-M2 (C-statistics = 0.806) model showed better differentiation success in morbidity than other models. The ACS-NSQIP-M1 (C-statistics = 0.888) and M2 (C-statistics = 0.956) models showed better mortality discrimination than P-POSSUM-M1 (C-statistics = 0.776) model. The P-POSSUM-M2 (C-statistics = 0.948) model showed better mortality differentiation success than the ACS-NSQIP-M1 and P-POSSUM-M1 models. The use of ACS-SRC and P-POSSUM calculators for major hepatobiliary surgeries offers quantitative data to assess risks for both the surgeon and the patient. Integrating these calculators into preoperative evaluation practices can enhance decision-making processes for patients. The results of the statistical analyses indicated that the P-POSSUM-M2 model for morbidity and the ACS-NSQIP-M2 model for mortality exhibited superior overall performance.

Abbreviations: ACS-NSQIP = American College of Surgeons National Surgical Quality Improvement Program, GCS = Glasgow Coma Scale, HL = Hosmer–Lemeshow, ICC = Intraclass Correlation Coefficient, M = model, O/E = observed/expected, P-POSSUM = Portsmouth Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity, SRC = surgical risk calculator.

Keywords: ACS score, hepatobiliary surgery, P-POSSUM score, surgical risk calculator

1. Introduction

Postoperative complications and the risk of mortality can be influenced by many factors. Hepatobiliary surgeries often have higher rates of mortality and complications.[1,2] These outcomes can vary based on factors such as the hospital where the surgery is conducted, the level of experience of the surgeons, the patients existing health conditions, and the specific type of surgery being performed.[1,2]

Pancreaticoduodenectomy, pylorus-preserving pancreaticoduodenectomy, total pancreatectomy, distal pancreatectomy, right hepatectomy, left hepatectomy, partial segmentectomy, hepatic trisegmentectomy, and hepaticojejunostomy surgeries are classified as major hepatobiliary surgeries.[3,4] Complications that may arise after these surgeries include pneumonia, cardiac complications, surgical site infections, urinary tract infections, venous thromboembolism, renal failure, and sepsis. Understanding the risks of these postoperative complications

The authors have no funding and conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are not publicly available, but are available from the corresponding author on reasonable request.

[a] Department of General Surgery, Istanbul University Cerrahpasa – Cerrahpasa School of Medicine, Istanbul, Turkey, [b] Department of General Surgery, Izmir University of Economics, School of Medicine, Izmir, Turkey.

* Correspondence: Ergin Erginoz, Department of General Surgery, Istanbul University Cerrahpasa – Cerrahpasa School of Medicine, Prof Dr Murat Dilmener Hastanesi, Bakirkoy Cd, Yesilkoy, 34153 Istanbul, Turkey (e-mail: eerginoz@ku.edu.tr).

How to cite this article: Karabulut A, Umman V, Oral G, Erginoz E, Carkman MS. Assessing the effectiveness of ACS surgical risk calculator versus P-POSSUM in predicting mortality and morbidity for major hepatobiliary surgery: An observational study. Medicine 2024;103:28(e38973).

Received: 9 May 2024 / Received in final form: 26 June 2024 / Accepted: 27 June 2024

http://dx.doi.org/10.1097/MD.0000000000038973

and making a surgical decision accurately is important for both the physician and the patient.[4,5]

For general surgical procedures, a scoring system developed according to the American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP) database is a tool that can be adjusted based on a surgeon's clinical experience and intuition and has good predictive ability, and it is a relatively easy-to-apply tool. This matured and tested surgical risk assessment calculator (ACS-SRC) has been developed to allow preoperative risk assessment for common surgical procedures based on ACS-NSQIP data. It utilizes 21 objective preoperative variables and has been validated for open pancreatic and laparoscopic/open colorectal, gallbladder, and hernia surgeries.[6] P-POSSUM, the Physiological and Operative Severity Score for the Enumeration of Morbidity and Mortality (POSSUM), has been proposed as a scoring system adjusted for risk to allow direct comparison between observed and expected adverse outcome rates. Therefore, it is referred to as a surgeon-based scoring system. Portsmouth POSSUM (P-POSSUM) is a modification of the POSSUM scoring system, containing the same variables and rating system but incorporating a different equation that better fits the observed mortality rate.[7] Such targeted assessment and prediction allow clinicians to approach decision-making with an evidence-based understanding of an individual's risk level. This precision has the potential to systematically optimize outcomes for surgical interventions.[7]

The increasing experience in major hepatobiliary surgeries over the years, advancements in technology, increased accessibility, and number of diagnostic tests have facilitated the ability to foresee risks before surgeries. Despite these advancements in operative and postdiagnostics, major hepatobiliary surgeries are still associated with high morbidity and mortality rates.[2] We aim to investigate whether these 2 risk calculators accurately reflect actual mortality and morbidity outcomes when retrospectively assessed based on preoperative risk scores, to compare their predictive superiority against each other, and to assess their applicability to major hepatobiliary surgery.

2. Materials and methods

2.1. Study groups

The study included 300 cases of major hepatopancreaticobiliary surgery conducted at a single institution between August 2016 and December 2021. Among the major hepatopancreaticobiliary cases, pancreaticoduodenectomy, pylorus-preserving pancreaticoduodenectomy, total pancreatectomy, distal pancreatectomy, hepatectomy (right, left, partial, trisegmentectomy), and hepaticojejunostomy surgeries were included. Surgeries examined under minor surgeries such as cholecystectomy and metastasectomy were excluded from the study. Fourteen cases of major hepatopancreaticobiliary surgery were excluded from the study due to the unavailability of patient data, incomplete data, and the emergency nature of the performed surgery. The study was approved by the Ethics Committee of Istanbul University Cerrahpasa – Cerrahpasa School of Medicine (Date: September 13, 2022, No. E-83045809-604.01.01-474698).

2.2. ACS-SRC and P-POSSUM risk scores

The American College of Surgeons surgical risk calculator (ACS-SRC) provides surgeons and patients with estimated risks for 12 postoperative complications along with predicted length of stay, using 21 preoperative factors (including comprehensive procedures, demographic characteristics, and comorbidities). This risk calculator collects high-quality, standardized clinical data on preoperative risk factors and postoperative complications. It gathers reliable and validated data on patient demographics, laboratory results, comorbidities, and 30-day postoperative outcomes for patients undergoing a wide range of operations across the majority of surgical specialties.

The POSSUM score is an 18-variable system designed to assist in predicting morbidity and mortality in general surgery. This model utilizes scores related to 12 physiological and 6 operative variables and was developed to predict postoperative 30-day mortality and morbidity. Among these variables are age, cardiac function, respiratory function (based on the degree of shortness of breath), electrocardiogram findings, systolic blood pressure, pulse rate, hemoglobin level, white cell count, serum urea, sodium, potassium, and Glasgow Coma Scale score. Using these variables, it generates scores for morbidity percentage and mortality percentage. Since it has been shown that the POSSUM score overestimates mortality, the Portsmouth prediction equation has been applied to improve the accuracy of the score.

2.3. Statistical analysis

Each patient was routinely assigned a morbidity and mortality risk score that was converted into a probability model 1 (M1). The variables used in the original P-POSSUM model were then entered into a binary logistic regression, and the estimated probabilities (ranging from 0 to 1) were used – model 2 (M2). This exercise was repeated for ACS-SRC. The receiver operating characteristic (ROC) curve was used to measure the diagnostic accuracy of the test. For both ACS-SRC and P-POSSUM, 2 ROC curves were derived for both M1 and M2. An area under the curve (AUC) value below 0.6 indicated poor discriminatory ability, 0.6 to 0.8 moderate discriminatory ability, 0.8 to 0.9 good discriminatory ability, and a value above 0.9 signified excellent discriminatory ability.

The comparison between observed and expected morbidity and mortality rates for each model was assessed using the Hosmer–Lemeshow (HL) goodness-of-fit test (*P* value). When the HL test reached nonsignificance ($P > .05$), the model had better goodness-of-fit. The observed/expected (O/E) ratio was used to evaluate calibration and analyze the ratio between the predicted morbidity/mortality based on scores for the relevant points in the dataset and the actual morbidity/mortality. A score of <1.0 indicates overestimation, while a score of >1.0 indicates underestimation. The Brier score was calculated as the mean squared deviation between the estimated and observed risks and used as a measure of accuracy. The lower the score, the better the performance and accuracy of individual morbidity/mortality risk estimation. The Brier score is reported as a score between 0 and 1 and is calculated as the mean squared difference between the predicted probability and the observed outcome. A score of 0 indicates no difference between the predicted and observed outcomes, thus representing the best possible test. A score of 1 indicates that the test did not predict the outcome.

3. Results

Among the 300 patients involved in the study, 158 (52.67%) were male and 142 (47.33%) were female. The mean age of the patients was 57.07 years (range: 18–88). The most commonly performed operation among major hepatobiliary surgeries was pylorus-preserving pancreaticoduodenectomy, with 66 cases (22.00%), followed by distal pancreatectomy with 65 cases (21.67%) (Table 1). The preoperative data that were entered into the ACS calculator is shown in Table 2 and the P-POSSUM preoperative data is shown in Table 3.

All patients underwent major surgery. The amount of bleeding during surgeries was as follows: in 17 patients (5.67%), <100 mL; in 139 patients (46.33%), between 101 and 500 mL; in 102 patients (34.00%), between 501 and 999 mL; and in 42 patients (14.00%), it was determined to be more than 1000 mL. Regarding contamination of the operative field, in 8 patients (2.67%) no contamination was observed. In 263 patients

(87.67%), minor contamination with free bowel content, pus, and blood was observed. In the study, 207 patients (69.00%) had cancer with lymph node metastasis, 28 patients (9.33%) had cancer with distant metastasis, 18 patients (6.00%) had only primary cancer without any nodal involvement, and 47 patients (15.67%) did not have cancer. All patients underwent elective surgery.

The postoperative findings from the ACS surgical risk calculator are shown in Table 4.

The mean length of hospital stay for patients was found to be 21.02 ± 11.43 days (range: 2–65 days). The estimated mean length of hospital stay calculated using the ACS Surgical Risk Calculator was 9.63 ± 5.25 days. According to the Intraclass Correlation Coefficient (ICC), which ranges from 0.00 to 0.50 indicating weak, 0.51 to 0.75 indicating moderate, 0.76 to 0.90 indicating good, and 0.91 to 1.00 indicating excellent results, the ICC was calculated as 0.447, indicating a weak result.

The coefficients of the variables that constitute the M2 models for morbidity for ACS-NSQIP and P-POSSUM have been calculated. Due to an insufficient number of observations in the fully dependent category of the functional status variable, it was not included in the ACS-NSQIP M2 model. Since there were no patients with ASA scores of 1, 4, and 5, this category was not used in the ACS-NSQIP M2 model. In the P-POSSUM M2 model, the Glasgow Coma Scale (GCS), type of operation, number of different surgeries to be performed, and urgency variables were not included. This is because these variables consist of a single category and their inclusion in the model would be meaningless.

Both the ACS-NSQIP and P-POSSUM M1 and M2 models' discriminatory ability, calibration, and overall performance for morbidity are shown in Table 5. The ACS-NSQIP M1 (C-statistic = 0.725) and M2 (C-statistic = 0.791) models demonstrated better discriminatory ability for morbidity compared to the P-POSSUM M1 model (C-statistic = 0.672). The P-POSSUM M2 model (C-statistic = 0.806) showed better discriminatory ability than the other models. Overall, according

Table 1

Type of operations performed that were included in the study.

Operation	n	(%)
Pancreaticoduodenectomy	52	(17.33)
Pylorus-preserving pancreaticoduodenectomy	66	(22.00)
Total pancreatectomy	26	(8.67)
Distal pancreatectomy	65	(21.67)
Right hepatectomy	5	(1.67)
Left hepatectomy	7	(2.33)
Trisegmentectomy	6	(2.00)
Partial segmentectomy	32	(10.67)
Hepaticojejunostomy	41	(13.67)

Table 2

The ACS surgical risk calculator preoperative parameters.

Parameters		n/mean ± SD	(%)/Median (Range)
Age		57,07 ± 13,87	59 (18–88)
Diabetes	Absent	223	(74.33)
	Insulin use	20	(6.67)
	Oral antidiabetic	57	(19.00)
Sex	Male	158	(52.67)
	Female	142	(47.33)
HT	Absent	174	(58.00)
	Present	126	(42.00)
Functional status	Independent	264	(88.00)
	Partially dependent	35	(11.67)
	Totally dependent	1	(.33)
Congestive heart failure	Absent	260	(86.67)
	Present	40	(13,33)
Dispnea	Absent	208	(69.33)
	At rest	3	(1.00)
	Medium level with effort	89	(29.67)
ASA Class	2	92	(30.67)
	3	208	(69.33)
Smoking	Absent	197	(65.67)
	Present	103	(34.33)
Steroid use	Absent	291	(97.00)
	Present	9	(3.00)
COPD	Absent	280	(93.33)
	Present	20	(6.67)
Ascites	Absent	285	(95.00)
	Present	15	(5.00)
Dialysis	Absent	294	(98.00)
	Present	6	(2.00)
Preoperative sepsis	Absent	283	(94.33)
	Present	17	(5.67)
Acute renal failure	Absent	298	(99.33)
	Present	2	(.67)
Ventilator dependent	Absent	213	(71.00)
	Present	87	(29.00)
Disseminated cancer	Absent	213	(71.00)
	Present	87	(29.00)
Height		170.38 ± 10.08	173 (150–196)
Weight		88.25 ± 12.58	88 (50–120)
BMI		30.51 ± 4.42	30.51 (20.2–44.44)

to the O/E ratio and Brier score, the P-POSSUM M1 model performed worse than the other models, while the P-POSSUM M2 model showed better overall performance than the other models.

In the subgroup analysis, the ACS-NSQIP M2 model (C-statistic = 0.809) outperformed the P-POSSUM M2 model (C-statistic = 0.790) in predicting morbidity in pancreatic surgeries, and overall, ACS-NSQIP performed better. In liver surgeries, the ACS-NSQIP (C-statistic = 0.865) was found to be superior to P-POSSUM M2 in predicting morbidity. However, in biliary surgeries, the P-POSSUM M2 (C-statistic = 0.825) was found to be superior to ACS-NSQIP M2 (C-statistic = 0.695) in predicting morbidity. Morbidities in liver surgeries showed weak goodness-of-fit in both models according to the HL test (*P* < .05).

In terms of mortality, the ACS-NSQIP M1 (C-statistic = 0.888) and M2 (C-statistic = 0.956) models demonstrated better discriminatory ability for mortality than the P-POSSUM M1 model (C-statistic = 0.776). The P-POSSUM M2 model (C-statistic = 0.948) showed better discriminatory ability for mortality than the ACS-NSQIP M1 and P-POSSUM M1 models. Overall, according to the observed to expected ratio and Brier score, the P-POSSUM M1 model performed worse than the other models, while the ACS-NSQIP M2 model showed better overall performance than the other models. The ACS-NSQIP M2 model (C-statistic = 0.945) also demonstrated superior performance in predicting mortality for pancreatic surgeries compared to the P-POSSUM M2 model (C-statistic = 0.937), and it also showed better overall performance. In terms of biliary surgeries, the ACS-NSQIP M2 model (C-statistic = 1.000)

Table 3

The P-POSSUM physiological parameters.

Parameters		n/meant ± SD	(%)/Median (Range)
Age		57.07 ± 13.87	59 (18–88)
Cardiac	Absent	172	(57.33)
	Angina requiring medical treatment or HT	126	(42.00)
	Peripheral edema, warfarin, borderline cardiomyopathy	1	(.33)
	Increased CVP, cardiomegaly	1	(.33)
Respiratory	No dyspnea	208	(69.33)
	Dyspnea with exercise, mild COPD	72	(24.00)
	Limiting dyspnea, medium COPD	17	(5.66)
	Resting dyspnea, fibrosis, or consolidation on X-ray	3	(1.00)
ECG	Normal	248	(82.67)
	AF, pulse between 60 and 90	8	(2.67)
	Other rhythm abnormalities and abnormalities	44	(14.67)
Systolic blood pressure		128.82 ± 14.05	126 (100–165)
Pulse Rate		80.02 ± 11.34	78 (54–130)
Hemoglobin		12.29 ± 1.88	12.3 (8.8–18.7)
WBC		7.78 ± 2.85	7.5 (2.3–31)
Ure		30.73 ± 13.78	28 (8–123)
Sodium		139.64 ± 4.45	140 (126–158)
Potassium		4.3 ±. 50	4.3 (2.6–5.9)
GCS		15 ±. 00	15 (15–15)

Table 4

The postoperative information from the ACS surgical risk calculator.

Complications	Absent n (%)	Present n (%)	Median risk % (Range)	Brier score
Severe complication	199 (66.33)	101 (33.67)	24.9 (5.3–73.4)	.0915
Pneumonia	254 (84.67)	46 (15.33)	3.8 (0.3–30.5)	.0018
Cardiac complication	226 (75.33)	74 (24.67)	1.65 (0–18.9)	.0004
Surgical site infection	184 (61.33)	116 (38.67)	15.95 (0.1–34.8)	.0473
Urinary tract infection	288 (96)	12 (4)	2.2 (0.6–20.8)	.0005
Venous thromboembolism	292 (97.33)	8 (2.67)	3.6 (0.5–37.9)	.0013
Renal failure	291 (97)	9 (3)	1.5 (0–58.4)	.0002
Readmission	232 (77.33)	68 (22.67)	15.8 (0.1–80.7)	.0262
Reoperation	285 (95)	15 (5)	5 (1.2–35.1)	.0025
Palliative care	120 (40)	180 (60)	6.6 (0.3–91.6)	.3064
Sepsis	275 (91.67)	25 (8.33)	8.1 (0–52.7)	.0072

Table 5

The morbidity comparison between ACS-NSQIP and P-POSSUM.

Model	C-statistics [95% CI]	Observed/expected ratio	Hosmer–Lemeshow(*P* value)	Brier score
ACS-NSQIP M1	0.725 [0.668–0.782]	1.36	.186	0.2156
ACS-NSQIP M2	0.791 [0.741–0.842]	1.33	.053	0.1958
POSSUM M1	0.672 [0.611–0.732]	1,48	.352	0.2567
POSSUM M2	0.806 [0.758–0.853]	1.21	.166	0.1327

outperformed the P-POSSUM M2 model (C-statistic = 0.975) in predicting mortality. In terms of the HL test for Pancreatic Surgeries, the P-POSSUM M2 model ($P = .004$) exhibited poor goodness-of-fit compared to the ACS-NSQIP M2 model ($P = .579$).

4. Discussion

Some reasons for choosing ACS-SRC and P-POSSUM in our study against other currently used risk calculators include their user-friendly interfaces, reliance on standardized data, evaluation of parameter variables routinely used in preoperative, intraoperative, and postoperative assessment, and the demonstrated inadequacy of some other risk calculators in predicting morbidity in hepatobiliary surgery.[4,6,7]

Surgeons not only want to inform patients but also want to know the surgical risks that may occur postoperatively. It is very important to determine changeable risk factors to improve the patient's condition before surgery and patient outcomes in the postoperative period. Surgical risk calculators have been developed to guide both doctor and patient in the preoperative decision-making process and to predict surgical outcomes.[8]

Because POSSUM overestimated mortality, Prytherch et al[9] changed this calculator and found P-POSSUM. Therefore, a modified POSSUM variant, the P-POSSUM calculator, was used in our study. P-POSSUM has been tested as a predictor of surgical outcomes in various branches of surgery, and most studies have shown it to be a valid risk calculator in predicting mortality. Its accuracy in predicting morbidity has not been adequately tested and it has been suggested that it gives variable results.[10]

Risk calculators such as ACS-NSQIP and P-POSSUM are derived using retrospective data collected preoperatively and on intraoperative factors. The clinical utility of a model depends not only on its statistical performance but also on other factors. The data used should be parameters generally evaluated before surgery. It is important to have a model that includes inexpensive laboratory tests, is designed to be simple for the user, and includes surgery-related variables. Hemodynamic parameters are used in P-POSSUM, and these may change due to external factors that do not depend on the patient, which may affect the mortality and morbidity risk results that may occur after surgery. ACS-NSQIP, on the other hand, uses clear data to determine risk and differs from P-POSSUM.[6,7]

Another notable point is that while the ACS-NSQIP surgical risk calculator is a preoperative risk prediction model in its structure, P-POSSUM is a perioperative model.[6,7] Since P-POSSUM requires preoperative data, it helps postoperatively rather than preoperatively. ACS-NSQIP will be more useful for preoperative decisions. The unique abilities of ACS-NSQIP and P-POSSUM may suggest that instead of competing, they could be used together to help clinicians make important perioperative decisions to improve postoperative outcomes.

In our study, when looking at major hepatobiliary surgeries, the ACS-NSQIP M1 and M2 models showed better discriminative ability for morbidity compared to the P-POSSUM M1 model. The P-POSSUM M2 model, on the other hand, demonstrated better discriminative ability than other models. Overall, while the P-POSSUM M1 model showed worse performance compared to other models based on the O/E ratio and Brier score, the P-POSSUM M2 model exhibited better overall performance than other models.

On the other hand, the ACS-NSQIP M1 and M2 models showed better discriminative ability for mortality compared to the P-POSSUM M1 model. The P-POSSUM M2 model, however, demonstrated better discriminative ability for mortality than the ACS-NSQIP M1 and P-POSSUM M1 models. Overall, while the P-POSSUM M1 model showed worse performance compared to other models based on the O/E ratio and Brier score, the ACS-NSQIP M2 model exhibited better overall performance than other models.

In a study conducted by Sudharasan et al in 2018, which included 245 patients who underwent liver resection, they demonstrated the superiority of ACS over POSSUM. The ACS-NSQIP score indicated better discrimination ability than POSSUM ($P = .03$). However, in the O/E ratio, it was determined that the ACS-NSQIP M1 model overestimated mortality, while the POSSUM M1 model underestimated mortality.[4]

In our study, subgroup analysis of major hepatobiliary surgeries was conducted for pancreatic, liver, and bile duct surgeries. The ACS-NSQIP M2 was found to be superior to P-POSSUM in predicting morbidity in pancreatic surgeries and also showed better overall performance. Similarly, for liver surgeries, ACS-NSQIP was better than P-POSSUM in estimating morbidity. However, P-POSSUM was found to be superior to ACS-NSQIP in predicting morbidity in bile duct surgeries. Morbidities in liver surgeries showed poor goodness-of-fit in terms of the HL test for both models.

ACS-NSQIP was found to be superior to P-POSSUM in predicting morbidity after major hepatobiliary surgeries. If we look at the reasons for this, choosing which surgery will be performed before calculating the risk in the ACS-NSQIP risk calculator makes it more specific. While P-POSSUM aggregates all major hepatobiliary surgeries under a single heading, ACS-NSQIP offers different options based on the surgical types. An example of this could be distinguishing between a partial hepatectomy and a right hepatectomy, which has been shown to carry a higher postoperative risk.[11,12] In a study led by Seung Jae Lee et al in 2020, the C-statistic value of ACS-NSQIP was found to be relatively low (0.626), and the HL goodness-of-fit test indicated weak calibration.[13] In our study, the C-statistic values in our ACS-NSQIP M2 model were found to be high, such as 0.791 for morbidity and 0.956 for mortality, and it was shown to be strongly calibrated for the HL goodness-of-fit test (>0.05).

Another limiting factor is that the ACS-NSQIP surgery calculator only involves 1-month after surgery results.[6] Including late complications could assess surgical risk more accurately. Especially in pancreatic surgeries, complications may occur in the late period, and these may include problems with anastomosis. In a cohort study conducted by Seung Jae Lee et al in 2020, it was estimated that severe late complications accounted for 39.0% of overall major complications.[13] Much work is needed to develop pancreatectomy-specific risk models that involve late complications.

The ACS-NSQIP and P-POSSUM databases do not collect disease-specific and procedure-specific data, and therefore information on various complications and outcome measures specific to hepatobiliary surgery, such as bile leakage, pancreatic fistula formation, functional status of the gastrointestinal system, liver failure, and portal vein thrombosis, is not available.[5,14]

5. Conclusions

The ACS-NSQIP M1 and M2 models showed better discriminative ability for morbidity compared to the P-POSSUM M1 model. The P-POSSUM M2 model, on the other hand, exhibited better discriminative performance than the other models. For mortality, the ACS-NSQIP M2 model demonstrated better overall performance for major hepatobiliary surgeries compared to other models. It is recommended that centers customize and improve the morbidity and mortality risk calculation models by adding parameters specific to the surgical subgroup to achieve more accurate results in major hepatobiliary surgeries.

Author contributions

Conceptualization: Ali Karabulut, Gunes Oral, Ergin Erginoz, Mehmet Sinan Carkman.
Resources: Ali Karabulut.
Validation: Ali Karabulut, Veysel Umman, Mehmet Sinan Carkman.
Writing – review & editing: Ali Karabulut, Mehmet Sinan Carkman.
Investigation: Veysel Umman.
Visualization: Veysel Umman.
Data curation: Gunes Oral, Ergin Erginoz.
Writing – original draft: Gunes Oral, Ergin Erginoz.
Formal analysis: Ergin Erginoz.
Supervision: Mehmet Sinan Carkman.

References

[1] Kneuertz PJ, Pitt HA, Bilimoria KY, et al. Risk of morbidity and mortality following hepato-pancreato-biliary surgery. J Gastrointest Surg. 2012;16:1727–35.

[2] Wang H, Chen T, Wang H, et al. 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. Am J Surg. 2013;205:466–72.

[3] Chincarini M, Zamboni GA, Pozzi Mucelli R. Major pancreatic resections: normal postoperative findings and complications. Insights Imaging. 2018;9:173–87.

[4] Madhavan S, Shelat VG, Soong S-L, et al. Predicting morbidity of liver resection. Langenbecks Arch Surg. 2018;403:359–69.

[5] Cusworth BM, Krasnick BA, Nywening TM, et al. Whipple-specific complications result in prolonged length of stay not accounted for in ACS-NSQIP Surgical Risk Calculator. HPB (Oxford). 2017;19:147–53.

[6] 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;217:833–42.e1.

[7] Tyagi A, Nagpal N, Sidhu DS, et al. Portsmouth physiological and operative severity score for the Enumeration of Mortality and morbidity scoring system in general surgical practice and identifying risk factors for poor outcome. J Nat Sci Biol Med. 2017;8:22–5.

[8] Rix TE, Bates T. Pre-operative risk scores for the prediction of outcome in elderly people who require emergency surgery. World J Emerg Surg. 2007;2:16.

[9] 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.

[10] Mukherjee S, Kedia A, Goswami J, et al. Validity of P-POSSUM in adult cancer surgery (PACS). J Anaesthesiol Clin Pharmacol. 2022;38:61–5.

[11] Helling TS, Blondeau B. Anatomic segmental resection compared to major hepatectomy in the treatment of liver neoplasms. HPB (Oxford). 2005;7:222–5.

[12] Karanjia ND, Lordan JT, Quiney N, et al. A comparison of right and extended right hepatectomy with all other hepatic resections for colorectal liver metastases: a ten-year study. Eur J Surg Oncol. 2009;35:65–70.

[13] Lee SJ, Park SY, Hwang DW, et al. Surgical decisions based on a balance between malignancy probability and surgical risk in patients with branch and mixed-type intraductal papillary mucinous neoplasm. J Clin Med. 2020;9:2758.

[14] Leichtle SW, Kaoutzanis C, Mouawad NJ, et al. Classic Whipple versus pylorus-preserving pancreaticoduodenectomy in the ACS NSQIP. J Surg Res. 2013;183:170–6.