ORIGINAL ARTICLE

Use of Machine Learning for Prediction of Patient Risk of Postoperative Complications After Liver, Pancreatic, and Colorectal Surgery

Katiuscha Merath[1] · J. Madison Hyer[1] · Rittal Mehta[1] · Ayesha Farooq[1] · Fabio Bagante[1] · Kota Sahara[1] · Diamantis I. Tsilimigras[1] · Eliza Beal[1] · Anghela Z. Paredes[1] · Lu Wu[1] · Aslam Ejaz[1] · Timothy M. Pawlik[1]

Abstract

Background Surgical resection is the only potentially curative treatment for patients with colorectal, liver, and pancreatic cancers. Although these procedures are performed with low mortality, rates of complications remain relatively high following hepatopancreatic and colorectal surgery.

Methods The American College of Surgeons (ACS) National Surgical Quality Improvement Program was utilized to identify patients undergoing liver, pancreatic and colorectal surgery from 2014 to 2016. Decision tree models were utilized to predict the occurrence of any complication, as well as specific complications. To assess the variability of the performance of the classification trees, bootstrapping was performed on 50% of the sample.

Results Algorithms were derived from a total of 15,657 patients who met inclusion criteria. The algorithm had a good predictive ability for the occurrence of any complication, with a C-statistic of 0.74, outperforming the ASA (C-statistic 0.58) and ACS-Surgical Risk Calculator (C-statistic 0.71). The algorithm was able to predict with high accuracy thirteen out of the seventeen complications analyzed. The best performance was in the prediction of stroke (C-statistic 0.98), followed by wound dehiscence, cardiac arrest, and progressive renal failure (all C-statistic 0.96). The algorithm had a good predictive ability for superficial SSI (C-statistic 0.76), organ space SSI (C-statistic 0.76), sepsis (C-statistic 0.79), and bleeding requiring transfusion (C-statistic 0.79).

Conclusion Machine learning was used to develop an algorithm that accurately predicted patient risk of developing complications following liver, pancreatic, or colorectal surgery. The algorithm had very good predictive ability to predict specific complications and demonstrated superiority over other established methods.

Keywords Complications · Machine learning · Liver · Pancreas · Colorectal

Introduction

Surgical resection is the only potentially curative treatment for patients with colorectal, liver, and pancreatic cancers. Although these procedures are now performed with very low mortality, the incidence of complications remains high following hepatopancreatic (HP) or colorectal surgery[123] Postoperative complications are deleterious events for pa-

tients, impacting perioperative mortality, cancer recurrence, patient-reported experiences, and hospital costs[4,5,6,7] For example, Silber et al. demonstrated that the development of a complication after surgery, even seemingly mild ones, may markedly alter the patient's long-term prognosis, and risk of death.[8] Among patients with cancer, the development of postoperative complications has been independently associated with worse recurrence-free survival (RFS) and overall survival (OS)[9,10] In addition, our group has demonstrated that complications interact in a synergistic manner among patients who develop more than one complication, increasing the risk of death exponentially, rather than in an additive way.[11] Moreover, postoperative complications result in increased length of hospital stay (LOS), which translates into increased hospital costs and a higher financial burden for patients[12,13,14]

✉ Timothy M. Pawlik
tim.pawlik@osumc.edu

[1] Department of Surgery, Division of Surgical Oncology, The Ohio State University Wexner Medical Center, 395 W. 12th Ave., Suite 670, Columbus, OH, USA

Accurately predicting the risk of a postoperative complication is important for appropriate patient selection prior to surgery, for guiding perioperative decision-making, determining the necessary level of vigilance in the postoperative period, as well as directing early interventions. A number of risk stratification and predictive tools have been proposed to stratify a patient's risk of postoperative morbidity and mortality, such as the American College of Surgeons Surgical Risk Calculator (ACS-SRC), American Society of Anesthesiologists (ASA) score, and the Physiologic and Operative Severity Score for the Enumeration of Mortality and Morbidity (POSSUM) score[15,16,17] Nevertheless, these predictive tools are subjective and have only a moderate predictive ability with limited clinical applicability[18,19,20] In addition, these predictive tools and risk calculators were developed based on the premise that the variables in the models interact in a linear and additive fashion. In clinical reality, however, the interaction between comorbidities and physiological factors may not be exactly linear.[21] Instead, a patient's risk for developing complications after a surgical procedure is multifactorial. Specifically, preoperative general health, physiologic capacity to withstand the surgical insult, type of anesthesia, and type of surgery all need to be factored into the equation. In turn, certain variables in a risk prediction model might gain or lose significance depending on the absence or presence of other factors.

Machine-learning techniques have been gaining popularity in the field of medicine as a more comprehensive, "non-linear", and accurate method to predict patient outcomes.[22] In particular, decision-tree algorithms, which sort through a vast number of variables looking for combinations that reliably predict outcomes, may be superior to the classic "linear" predictive tools currently used by clinicians for patient prognostication[23,24] To this end, the objective of the current study was to develop a machine-learning algorithm to predict the risk of postoperative complications following liver, pancreatic, and colorectal surgery using a large, national database.

Methods

Data Sources and Study Population

The American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) database is the largest, most reliable, and best-validated database in surgery[25,26,27] Using the ACS-NSQIP participant use data file (PUF), patients who underwent hepatic, pancreatic, and colorectal surgery between 2014 and 2016 were identified. Information in the final dataset included preoperative comorbidities and perioperative clinical variables, as well as 30-day postoperative complications and mortality. Patients who underwent emergency surgery and individuals who were younger than 18 years of age were excluded. The analyses were performed on complete cases only; patients with missing data on the variables used to develop the algorithm were excluded.

Development of the Machine Learning-Based Decision-Tree Learning Algorithm and Statistical Analyses

Variables associated with preoperative patient characteristics were used to design the models, while variables representing postoperative complications were considered as dependent variables. The seventeen outcomes of interest included incidence of superficial surgical site infection (SSI), deep incisional SSI, pulmonary embolism, organ space SSI, sepsis, wound dehiscence, urinary tract infection, deep vein thrombosis, myocardial infarction, pneumonia, unplanned intubation, stroke, cardiac arrest, septic shock, bleeding requiring transfusion, progressive renal insufficiency, and use of ventilator for > 48 h.

Classification trees were constructed to create decision tree models to predict the occurrence of any complication in addition to the occurrence of specific complications (Appendix 1).[28] Classification tree learning is a common machine learning approach to classify patients into distinct groups with distinct outcomes. The algorithm iteratively dichotomizes patients until an end node is reached. Variables used to construct the tree may be used more than once and variables closer to the top of the tree (also known as "root node") tend to have more clinical significance than those closer to the end nodes (also known as "leaves").

The performance of the decision-tree learning algorithm to predict any 30-day postoperative complication, as well as each of the postoperative complications was measured by the C-statistic, also known as the area under the curve (AUC) and was obtained utilizing logistic regression.[29] The C-statistic is a measure of concordance between model-based risk estimates and observed events, which has been used as a measure of model performance in several prior risk prediction models.[30] The classification tree grouped patients into distinct end nodes and provided an estimate for the probability of a complication. As such, each patient was assigned a probability of complication, and that value was used as the sole predictor in a logistic regression to assess the predictive performance of the algorithm. To assess the variability of the performance of the classification trees, bootstrapping was performed on 50% samples.[31] Bootstrapping is a statistical method in which multiple subsamples of the original population are taken, and statistical analyses are applied to each subsample in an effort to more accurately assess the variability of parameter estimates. To compare the performance of the current algorithm to the ASC-SRC, the prediction of morbidity included in the ACS-NSQIP data under the MORBPROB field was utilized.

Descriptive statistics for baseline characteristics were presented as median (interquartile range [IQR]) and frequency

(%) for continuous and categorical variables, respectively. To assess differences among baseline characteristics relative to the type of surgical procedure, Kruskal-Wallis one-way analysis of variance and chi-squared tests were used for continuous and categorical variables, respectively. All analyses were preformed using SAS v9.4.

Results

The comprehensive decision-tree learning algorithms were derived out of a total of 15,657 patients undergoing hepatopancreatic (HP) and colorectal surgery who met inclusion criteria. In total, 685 patients who underwent hepatic surgery, 6012 who underwent pancreatic surgery, and 8960 who underwent colorectal surgery were included. The baseline characteristics of patient cohort utilized for the algorithm development are described in Table 1. Among patients included in the study who had a liver resection, 54.9% ($n = 376$) had a postoperative complication, whereas 52.3% ($n = 3142$) and 28.5% ($n = 2555$) of patients who underwent a pancreatic and colorectal surgery, respectively, developed a complication following surgery ($p < 0.001$).

Thirty-Day Postoperative Morbidity and 30-Day Individual Postoperative Complications

The decision-tree algorithm for the occurrence of any complication within 30 days following a liver, pancreatic, or colorectal resection is depicted in Fig. 1. The algorithm had good predictive ability for the occurrence of any complication, with a c-statistic of 0.74, outperforming the ASA (c-statistic 0.58) and providing similarly strong results as the ACS-SRC (c-statistic 0.71).

The algorithm predicted the occurrence of specific 30-day postoperative complications with a good to extremely high accuracy (c-statistic range from 0.76 to 0.98) (Table 2). The algorithm was also able to predict with high accuracy (c-statistic > 0.8); thirteen out of the seventeen postoperative complications analyzed. The best performance was in the prediction of a stroke (c-statistic 0.98), followed by wound dehiscence, cardiac arrest, and progressive renal failure (all c-statistic 0.96). The algorithm had a good predictive ability (c-statistic > 0.7) for other outcomes such as superficial SSI (c-statistic 0.76), organ space SSI (c-statistic 0.76), sepsis (c-statistic 0.79), and bleeding requiring transfusion (c-statistic 0.79).

Fig. 1 depicts the decision-tree algorithm for the calculation of the risk to develop any complication after HP or colorectal surgery. Figure 2 depicts the top of the decision-tree that displays the root node (node zero). The first split in the tree is between the types of surgical procedure, which were categorized by the algorithm as HP and colorectal, highlighting the distinction between the procedure types and allowing for the remainder of the tree to be procedure-specific. Following procedure type, the next variable sorted by the algorithm was patient hematocrit. Subsequently, the decision-tree defined different "questions" depending on the response to the previous question, and the direction was determined by successive individual responses based on the machine learning algorithm. For example, in Fig. 2, the third variable analyzed by the algorithm was "albumin" for patients undergoing colorectal surgery with a hematocrit lower than 32.2 g/dl, while for patients undergoing colorectal surgery with a hematocrit higher than 32.2 g/dl the next variable analyzed was the surgical approach. Of note, the optimal cutoff values for each continuous variable introduced in the decision-tree were determined by the machine-based learning algorithm itself. Unique decision-tree nodes were developed for each patient as different values or responses were introduced, until a final individual risk was generated.

Discussion

Machine learning is a subfield of artificial intelligence comprising a wide variety of data-driven methods and algorithms that use historical data for acquiring knowledge and making predictions or inferences on new data[25,32] These algorithms have the ability to learn and improve its own performance with time, allowing predictions to become increasingly more accurate as the model is exposed to more information. The application of machine-learning methodologies is relatively new in medicine, and there remains a substantial unfamiliarity around its utility in healthcare. Among the numerous possible applications, machine-learning techniques can be applied to large clinical datasets for the development of robust risk models. To this point, the current study was important because it developed a machine-learning algorithm to predict the risk of postoperative complications following HP and colorectal surgery. The algorithm developed in the current study demonstrated a good ability to predict patient risk of developing any complication following HP and colorectal surgery, outperforming the ASC-SRC and the ASA score. In addition, the accuracy of the algorithm was extremely high for the vast majority of the specific complications analyzed, with the c-statistics ranging from 0.76 for superficial SSI and organ space SSI to 0.98 for stroke.

Although the practical application of machine learning techniques to the medical field has been relatively limited compared with other fields, in recent years an increasing number of studies have proposed the utilization of machine learning algorithms for different purposes. For example, Zhao et al. used machine learning to develop an accurate predictive model for case duration of robotic-assisted surgery in order to increase utilization of robotic units during block time.[33] In a

Table 1 Preoperative characteristics of patients undergoing colorectal, liver, and pancreatic surgery who met the inclusion criteria

Preoperative characteristics	Total $N = 15{,}657$	Liver $N = 685$	Pancreas $N = 6012$	Colorectal $N = 8960$	p value
Age (median, IQR)	66 (57, 75)	65 (57, 72)	66 (58, 73)	66 (56, 76)	< 0.001
ASA class					< 0.001
1	104 (0.7%)	2 (0.3%)	19 (0.3%)	83 (0.9%)	
2	3927 (25.1%)	158 (23.1%)	1274 (21.2%)	2495 (27.8%)	
3	10,412 (66.5%)	478 (69.8%)	4365 (72.6%)	5569 (62.2%)	
4	1210 (7.7%)	47 (6.9%)	354 (5.9%)	809 (9%)	
5	4 (0%)	0 (0.0%)	0 (0.0%)	4 (0%)	
Functional status					< 0.001
Independent	15,222 (97.2%)	682 (99.6%)	5957 (99.1%)	8583 (95.8%)	
Partially dependent	381 (2.4%)	3 (0.4%)	50 (0.8%)	328 (3.7%)	
Totally dependent	54 (0.3%)	5 (0.1%)	0 (0.0%)	49 (0.5%)	
Dyspnea					< 0.001
At rest	57 (0.4%)	2 (0.3%)	11 (0.2%)	44 (0.5%)	
Moderate exertion	1156 (7.4%)	32 (4.7%)	334 (5.6%)	790 (8.8%)	
No	14,444 (92.3%)	651 (95%)	5667 (94.3%)	8126 (90.7%)	
Clinical T stage					< 0.001
T1	1598 (10.2%)	180 (26.3%)	618 (10.3%)	800 (8.9%)	
T2	2603 (16.6%)	297 (43.4%)	920 (15.3%)	1386 (15.5%)	
T3	9212 (58.8%)	162 (23.6%)	4104 (68.3%)	4946 (55.2%)	
T4	2244 (14.3%)	46 (6.7%)	370 (6.2%)	1828 (20.4%)	
Clinical N stage					< 0.001
N0	7835 (50%)	489 (71.4%)	2460 (40.9%)	4886 (54.5%)	
N1	6385 (40.8%)	184 (26.9%)	3552 (59.1%)	2649 (29.6%)	
N2	1437 (9.2%)	12 (1.8%)	0 (0.0%)	1425 (15.9%)	
Diabetes	1542 (9.8%)	46 (6.7%)	804 (13.4%)	692 (7.7%)	< 0.001
Smoking	2443 (15.6%)	139 (20.3%)	1028 (17.1%)	1276 (14.2%)	< 0.001
COPD	813 (5.2%)	35 (5.1%)	255 (4.2%)	523 (5.8%)	< 0.001
Ascites	128 (0.8%)	5 (0.7%)	21 (0.3%)	102 (1.1%)	< 0.001
CHF	224 (1.4%)	4 (0.6%)	24 (0.4%)	196 (2.2%)	< 0.001
Hypertension	8648 (55.2%)	365 (53.3%)	3316 (55.2%)	4967 (55.4%)	0.54
Acute renal failure	39 (0.2%)	1 (0.1%)	8 (0.1%)	30 (0.3%)	0.045
CRF on dialysis	92 (0.6%)	1 (0.1%)	15 (0.2%)	76 (0.8%)	< 0.001
Chronic steroid use	505 (3.2%)	21 (3.1%)	179 (3%)	305 (3.4%)	0.34
> 10% loss body weight*	1788 (11.4%)	62 (9.1%)	1030 (17.1%)	696 (7.8%)	< 0.001
Bleeding disorder	720 (4.6%)	23 (3.4%)	205 (3.4%)	492 (5.5%)	
Pre-operative transfusion	548 (3.5%)	6 (0.9%)	64 (1.1%)	478 (5.3%)	< 0.001
Serum Na^+ (median, IQR)	139 (137, 141)	139 (137, 141)	139 (137, 141)	139 (137, 141)	< 0.001
BUN	14 (10, 18)	14 (11, 18)	14 (11, 18)	13 (10, 18)	< 0.001
Creatinine	0.83 (0.7, 1)	0.83 (0.7, 1)	0.8 (0.7, 1)	0.86 (0.7, 1.03)	< 0.001
Albumin	3.8 (3.4, 4.2)	3.9 (3.5, 4.2)	3.9 (3.4, 4.2)	3.8 (3.3, 4.2)	< 0.001
Total bilirubin	0.5 (0.38, 0.8)	0.6 (0.4, 1)	0.7 (0.4, 1.59)	0.5 (0.3, 0.7)	< 0.001
AST	22 (17, 33)	33 (24, 57)	28 (20, 49)	20 (15, 26)	< 0.001
Alkaline phosphatase	85 (66, 121)	111 (77, 194)	112 (77, 211.5)	76 (62, 95)	< 0.001
WBC	6.9 (5.55, 8.6)	6.92 (5.7, 8.7)	6.9 (5.6, 8.5)	6.9 (5.5, 8.7)	0.37
HCT	36.8 (32.3, 40.6)	38.7 (35, 42.1)	37.8 (34.1, 41)	35.6 (31, 40.1)	< 0.001
Platelet count	249 (198, 313)	237 (185, 304)	240 (190, 298)	257 (204, 325)	< 0.001
PTT	29.5 (26.9, 32.3)	30 (27, 32.9)	29.6 (26.9, 32.2)	29.4 (26.8, 32.2)	0.1
INR	1 (1, 1.1)	1 (1, 1.1)	1 (1, 1.1)	1.02 (1, 1.1)	< 0.001
PT	12 (10.9, 13.3)	12.3 (11, 13.6)	11.7 (10.8, 13)	12.3 (11.1, 13.6)	< 0.001

Table 1 (continued)

| Preoperative characteristics	Total $N = 15{,}657$	Liver $N = 685$	Pancreas $N = 6012$	Colorectal $N = 8960$	p value
Minimally invasive surgery	7090 (45.3%)	81 (11.8%)	1018 (16.9%)	5991 (66.9%)	< 0.001
30-day mortality	266 (1.7%)	35 (5.1%)	91 (1.5%)	140 (1.6%)	< 0.001
Postoperative complication	6073 (38.8%)	376 (54.9%)	3142 (52.3%)	2555 (28.5%)	< 0.001

*in the last 6 months; *COPD* chronic obstructive pulmonary disease; *CHF* congestive heart failure; *CRF* chronic renal failure; *BUN* blood urea nitrogen; *AST* aspartate aminotransferase; *WBC* white blood count; *HCT* hematocrit; *PTT* partial thromboplastin time; *INR* international normalized ratio; *PT* prothrombin time

separate study, Harvin et al. used machine learning to identify factors associated with changes in surgical decision-making regarding the use of damage control laparotomy following the implementation of a quality improvement intervention.[34] Moreover, Canchi et al. reviewed the use of machine learning models to predict the risk of rupture among patients with

Fig. 1 A macro view of the decision-tree algorithm for the prediction of patient risk of developing any complication following colorectal, liver, or pancreatic surgery is presented. Red nodes depict a population where a majority had a complication, and blue nodes represent a population where a minority had a complication
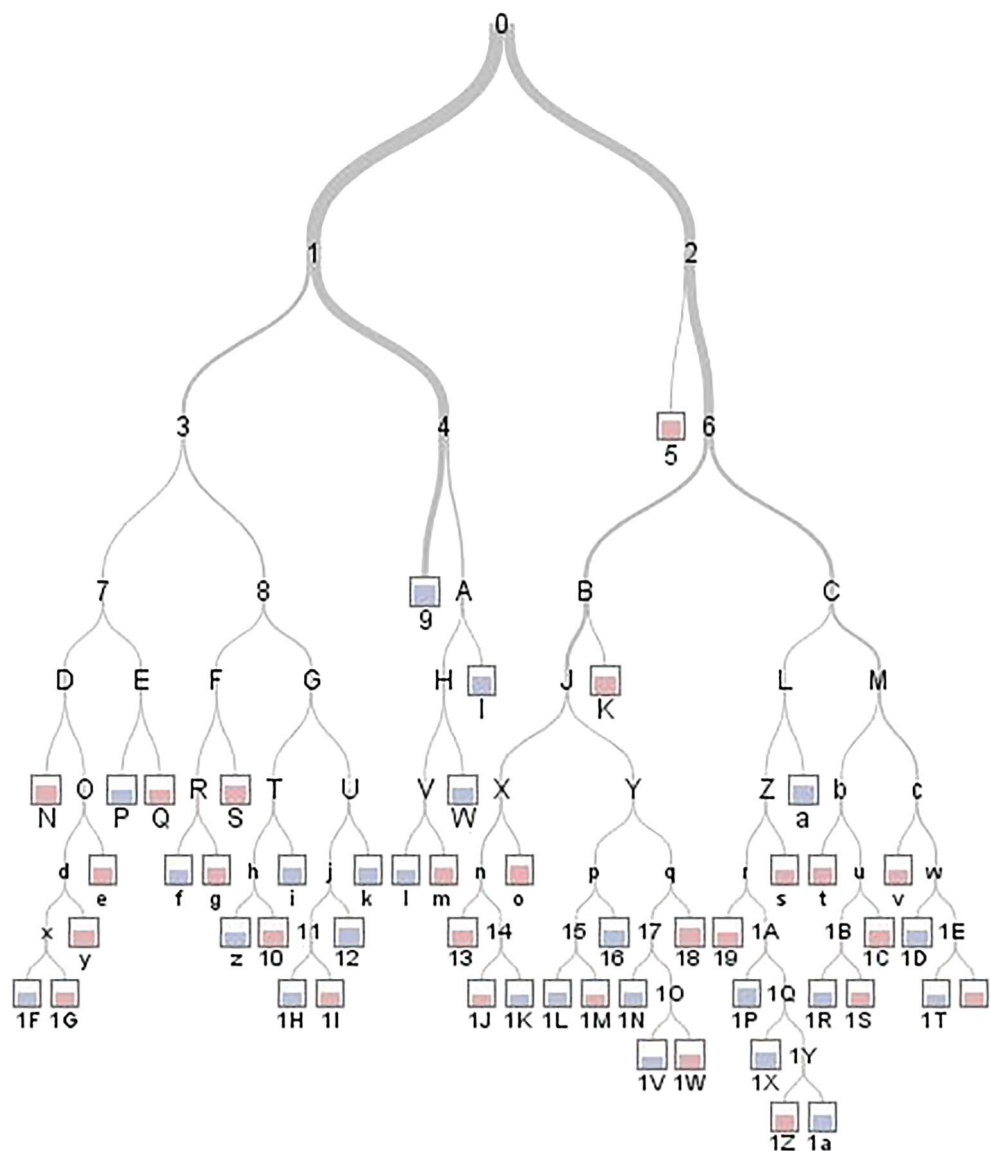
Table 2 Predictive ability of the algorithm for specific postoperative complications as measured by the C-statistic

| Outcome	C-statistic
Stroke	0.98 (0.97, 0.99)
Cardiac arrest	0.96 (0.92, 0.98)
Progressive renal insufficiency	0.96 (0.92, 0.97)
Wound dehiscence	0.96 (0.94, 0.98)
Myocardial infarction	0.95 (0.93, 0.97)
Deep incisional SSI	0.93 (0.9, 0.95)
Pulmonary embolism	0.93 (0.88, 0.95)
Ventilator > 48 h	0.88 (0.86, 0.9)
Septic shock	0.87 (0.85, 0.89)
DVT	0.87 (0.84, 0.89)
Unplanned intubation	0.86 (0.84, 0.88)
Pneumonia	0.84 (0.82, 0.86)
Urinary tract infection	0.82 (0.78, 0.85)
Sepsis	0.79 (0.78, 0.81)
Bleeding requiring transfusion	0.79 (0.77, 0.8)
Superficial surgical site infection	0.76 (0.74, 0.78)
Organ space SSI	0.76 (0.74, 0.77)
Any complication	0.74 (0.73, 0.75)

abdominal aortic aneurysms.[35] The authors noted that using machine learning to analyze complex radiological data combined with clinical data was a valuable tool for surgical decision-making.[35] Moreover, Yeong et al. used machine learning to develop a noninvasive method to predict skin burn healing time.[36] In fact, this machine-based model was able to predict which burns would heal in less than 14 days with 96% predictive accuracy.

Compared with conventional regression analyses, machine learning algorithms can improve predictive accuracy by capturing complex, nonlinear relationships observed in the data.[37] In the current study, a decision-tree learning approach was used to develop a risk model to predict complications after complex abdominal surgery. In our model, patient-level variables used to develop the decision-tree included a large number of clinical and laboratory parameters that are typically readily available in the patient electronic health record (EHR) (Table 1). The use of decision-tree learning methodology has the advantage of easy interpretability and high accuracy, as demonstrated in previous studies[38,39] With the use of decision-tree learning, the algorithm developed in the current study learned from actual patient outcomes as it was able to perform a comprehensive analysis of the interactions between all predictive factors influencing a patient's risk of developing a complication. In the process of estimating the risk of developing complications after surgery, the decision trees take different directions depending on the "answer" to each "question," with different variables being subsequently required

for the final risk prediction (Fig. 1). To this point, instead of adopting a linear method, the algorithm presented in the current study was able to detect the complex relationship among different variables and calculate individual patient risk in a more individualized manner.

The performance of the algorithm developed in the current study demonstrated superiority over established methods. Specifically, when we compared the accuracy of our decision-tree as a predictive model, the C-statistic for the prediction of patient risk of developing a complication following HP or colorectal surgery was 0.74, which performed similarly to the ASC-SRC (C-statistic 0.71) and outperformed the ASA score (C-statistic 0.58). In addition, when analyzing the risk of developing specific complications, the predictive ability was extremely high for the majority of the outcomes, ranging from 0.76 to 0.98 (Table 2). In the preoperative period, the accurate measurement of patient risk can not only facilitate a discussion about the risks and benefits of surgery during informed consent but also serves to identify patients who would benefit from preoperative strategies that could offset the risk. While it is true that many preoperative patient-level factors are non-modifiable, several measures can be adopted in order to mitigate patient risk of developing certain complications[40,41] Patients at increased risk of developing progressive renal insufficiency should not be prescribed nephrotoxic medications in the perioperative period. In the future, the integration of the current algorithm to patient EHRs will allow providers to make a timely and accurate prediction of patient risk of developing complications after surgery, so as to adopt tailored preventive measures and aid bedside decision-making.

Several limitations should be considered when interpreting the results of the current study. The performance and generalizability of machine-learning algorithms are dependent on the quality of data analyzed and, as with any retrospective study, selection biases resulting from the data collection methodology adopted by the ACS-NSQIP was a possibility. Moreover, the algorithm only included variables obtained by the ACS-NSQIP. To this point, it is possible that other preoperative factors not measured in the current algorithm will be more important predictive factors of patient risk. Nevertheless, the ACS-NSQIP database is widely recognized as the largest, most reliable, risk-adjusted and case-mix-adjusted, validated database in surgery[25,42] In addition, although certain outcomes of interest may extend beyond the 30 postoperative days, outcomes in the ACS-NSQIP data were limited to a 30-day follow-up period. The current study also did not assess patient risk of developing a procedure-specific complications such as anastomotic leak, pancreatic fistula, or postoperative liver failure. Furthermore, although the algorithm was able to estimate risk with high accuracy, causality between variables and outcomes cannot be determined
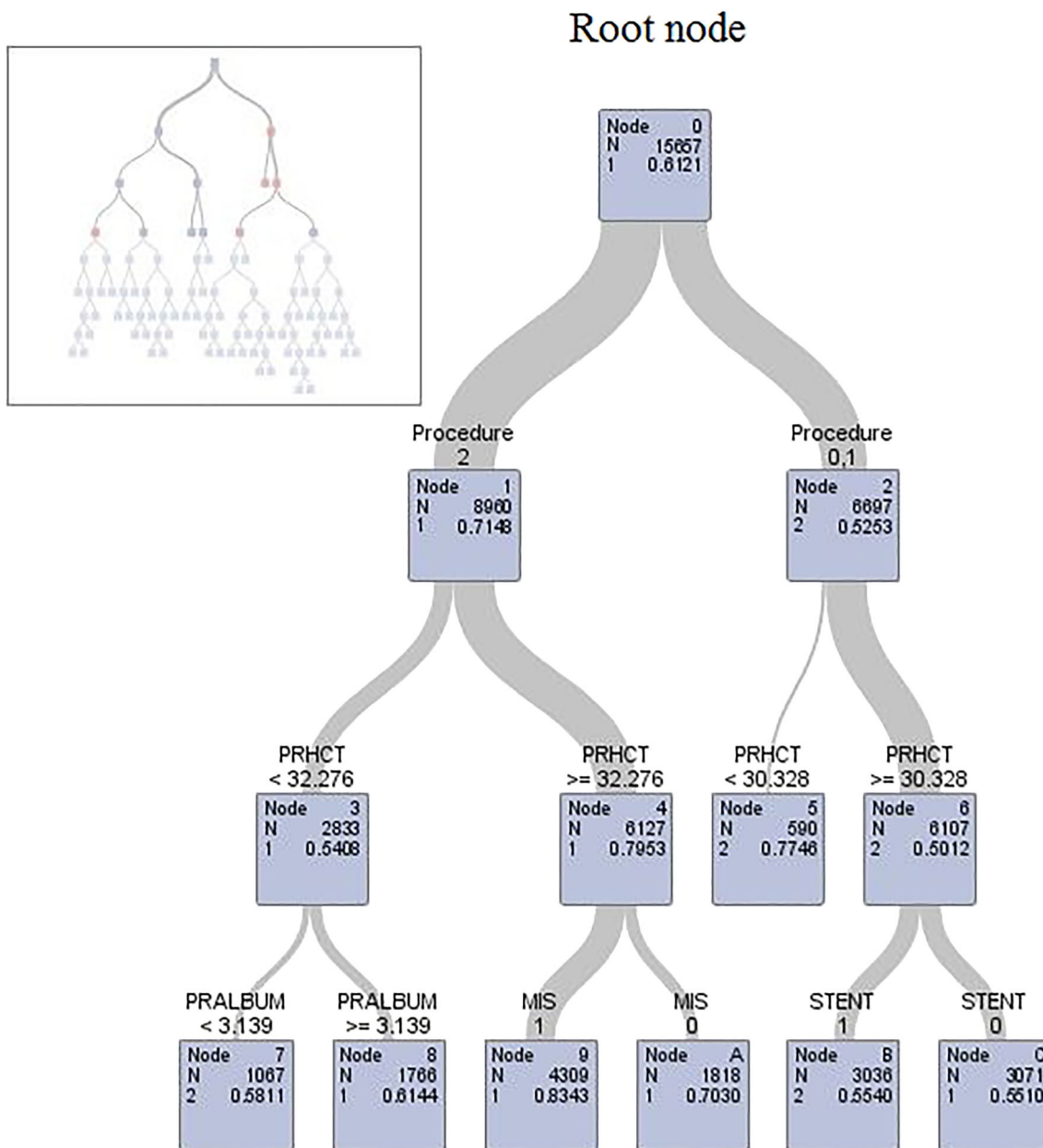
Fig. 2 A micro view of the decision-tree algorithm started at node 0 is presented for the prediction of patient risk of developing any complication following colorectal, liver, or pancreatic surgery. Each node of the presented subtree corresponds to a node in the macro view (Fig. 1) and provides details of which variables were first used to classify patients as well as what cutoffs were used for continuous variables

based solely on the decision tree. Finally, while we did internally validate the results of the classification tree, further research is needed to externally validate the results of the current algorithm.

In conclusion, we used machine learning to develop and validate a decision-tree learning–based algorithm that uses data readily available in patient EHRs for the prediction of patient risk of developing complications following liver, pancreatic, or colorectal surgery. The algorithm presented in the current study had good predictive ability as measured by the *C*-statistic and outperformed the ASA risk classification and the ACS-SRC. Future studies assessing the feasibility of the

integration of the algorithm to patient EHRs and its implementation to the clinical practice are warranted.

References

1. Merath K, Chen Q, Bagante F, Sun S, Akgul O, Idrees JJ, et al. Variation in the cost-of-rescue among medicare patients with complications following hepatopancreatic surgery. HPB (Oxford) [Internet]. 2018 Sep; Available from: https://www.ncbi.nlm.nih.gov/pubmed/30266495

2. Ghaferi AA, Birkmeyer JD, Dimick JB. Complications, failure to rescue, and mortality with major inpatient surgery in medicare patients. Ann Surg. 2009;250(6):1029–34.

3. Kohlnhofer BM, Tevis SE, Weber SM, Kennedy GD. Multiple complications and short length of stay are associated with postoperative readmissions. Am J Surg. 2014;207(4):449–56.

4. Mavros MN, de Jong M, Dogeas E, Hyder O, Pawlik TM. Impact of complications on long-term survival after resection of colorectal liver metastases. Br J Surg. 2013;100(5):711–8.

5. Healy MA, Mullard AJ, Campbell DA, Dimick JB. Hospital and payer costs associated with surgical complications. JAMA Surg. 2016;151(9):823–30.

6. Idrees JJ, Johnston FM, Canner JK, Dillhoff M, Schmidt C, Haut ER, et al. Cost of major complications after liver resection in the United States: are high-volume centers cost-effective? Ann Surg [Internet]. 2017; Available from: https://www.ncbi.nlm.nih.gov/pubmed/29232212

7. Tevis SE, Kennedy GD. Postoperative complications and implications on patient-centered outcomes. J Surg Res. 2013;181(1):106–13.

8. Silber JH, Rosenbaum PR, Trudeau ME, Chen W, Zhang X, Kelz RR, et al. Changes in prognosis after the first postoperative complication. Med Care. 2005;43(2):122–31.

9. Spolverato G, Yakoob MY, Kim Y, Alexandrescu S, Marques HP, Lamelas J, et al. Impact of complications on long-term survival after resection of intrahepatic cholangiocarcinoma. Cancer. 2015;121(16):2730–9.

10. Dorcaratto D, Mazzinari G, Fernandez M, Muñoz E, Garcés-Albir M, Ortega J, et al. Impact of postoperative complications on survival and recurrence after resection of colorectal liver metastases: systematic review and meta-analysis. Ann Surg. 2019;

11. Merath K, Chen Q, Bagante F, Akgul O, Idrees JJ, Dillhoff M, et al. Synergistic effects of perioperative complications on 30-day mortality following hepatopancreatic surgery. Journal of Gastrointestinal Surgery. 2018;22(10):1715–23.

12. Ejaz A, Kim Y, Spolverato G, Taylor R, Hundt J, Pawlik TM. Understanding drivers of hospital charge variation for episodes of care among patients undergoing hepatopancreatobiliary surgery. HPB (Oxford). 2015;17(11):955–63.

13. Abdelsattar ZM, Birkmeyer JD, Wong SL. Variation in Medicare payments for colorectal cancer surgery. J Oncol Pract. 2015;11(5):391–5.

14. Pradarelli JC, Healy MA, Osborne NH, Ghaferi AA, Dimick JB, Nathan H. Variation in Medicare expenditures for treating perioperative complications: the cost of rescue. JAMA Surg. 2016;151(12):e163340.

15. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg. 1991;78(3):355–60.

16. Gawande AA, Kwaan MR, Regenbogen SE, Lipsitz SA, Zinner MJ. An Apgar score for surgery. J Am Coll Surg. 2007;204(2):201–8.

17. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, 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(5):833–842.e1-3.

18. Beal EW, Saunders ND, Kearney JF, Lyon E, Wei L, Squires MH, et al. Accuracy of the ACS NSQIP online risk calculator depends on how you look at it: results from the United States Gastric Cancer Collaborative. Am Surg. 2018;84(3):358–64.

19. Gleeson EM, Shaikh MF, Shewokis PA, Clarke JR, Meyers WC, Pitt HA, et al. WHipple-ABACUS, a simple, validated risk score for 30-day mortality after pancreaticoduodenectomy developed using the ACS-NSQIP database. Surgery. 11;160(5):1279–87.

20. Beal EW, Lyon E, Kearney J, Wei L, Ethun CG, Black SM, et al. Evaluating the American College of Surgeons National Surgical Quality Improvement project risk calculator: results from the U.S.

Extrahepatic Biliary Malignancy Consortium. HPB (Oxford). 2017;19(12):1104–11.

21. 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. Annals of Surgery. 2018;268(4):574–83.

22. Carlos RC, Kahn CE, Halabi S. Data science: big data, machine learning, and artificial intelligence. J Am Coll Radiol. 2018;15(3 **Pt B**):497–8.

23. Linden A, Yarnold PR. Modeling time-to-event (survival) data using classification tree analysis. J Eval Clin Pract. 2017;23(6):1299–308.

24. Hu Y-J, Ku T-H, Jan R-H, Wang K, Tseng Y-C, Yang S-F. Decision tree-based learning to predict patient controlled analgesia consumption and readjustment. BMC Med Inform Decis Mak. 2012;12:131.

25. Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, et al. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg. 1998;228(4):491–507.

26. Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY. Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: an evaluation of all participating hospitals. Ann Surg. 2009;250(3):363–76.

27. Fink AS, Campbell DA, Mentzer RM, Henderson WG, Daley J, Bannister J, et al. The National Surgical Quality Improvement Program in non-veterans administration hospitals: initial demonstration of feasibility. Ann Surg. 2002;236(3):344–53; discussion 353–354.

28. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, NY: Springer; 2009. 745 p. (Springer series in statistics).

29. Pencina MJ, D'Agostino RB. Evaluating discrimination of risk prediction models: the C statistic. JAMA. 2015;314(10):1063–4.

30. Merath K, Bagante F, Beal EW, Lopez-Aguiar AG, Poultsides G, Makris E, et al. Nomogram predicting the risk of recurrence after curative-intent resection of primary non-metastatic gastrointestinal neuroendocrine tumors: an analysis of the U.S. Neuroendocrine Tumor Study Group. J Surg Oncol [Internet]. 2018; Available from: https://www.ncbi.nlm.nih.gov/pubmed/29448303

31. Efron B, Tibshirani R. An introduction to the bootstrap. New York: Chapman & Hall; 1993. 436 p. (Monographs on statistics and applied probability).

32. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med. 2016;375(13):1216–9.

33. Zhao B, Waterman RS, Urman RD, Gabriel RA. A machine learning approach to predicting case duration for robot-assisted surgery. J Med Syst. 2019;43(2):32.

34. Harvin JA, Green CE, Pedroza C, Tyson JE, Moore LJ, Wade CE, et al. Using machine learning to identify change in surgical decision making in current use of damage control laparotomy. J Am Coll Surg. 2019;228(3):255–64.

35. Canchi T, Kumar SD, Ng EYK, Narayanan S. A review of computational methods to predict the risk of rupture of abdominal aortic aneurysms. Biomed Res Int. 2015;2015:861627.

36. Yeong E-K, Hsiao T-C, Chiang HK, Lin C-W. Prediction of burn healing time using artificial neural networks and reflectance spectrometer. Burns. 2005;31(4):415–20.

37. Chen JH, Asch SM. Machine learning and prediction in medicine - beyond the peak of inflated expectations. N Engl J Med. 2017;376(26):2507–9.

38. James G, Witten D, Hastie T, Tibshirani R, editors. An introduction to statistical learning: with applications in R. New York: Springer; 2013. 426 p. (Springer texts in statistics).

39. Parikh SA, Gomez R, Thirugnanasambandam M, Chauhan SS, De Oliveira V, Muluk SC, et al. Decision tree-based classification of abdominal aortic aneurysms using geometry quantification measures. Ann Biomed Eng. 2018;46(12):2135–47.

40. Wynter-Blyth V, Moorthy K. Prehabilitation: preparing patients for surgery. BMJ. 8;358:j3702.

41. Minnella EM, Liberman AS, Charlebois P, Stein B, Scheede-Bergdahl C, Awasthi R, et al. The impact of improved functional capacity before surgery on postoperative complications: a study in colorectal cancer. Acta Oncol. 2019;1–6.

42. Raval MV, Pawlik TM. Practical guide to surgical data sets: national surgical quality improvement program (NSQIP) and pediatric NSQIP. JAMA Surg [Internet]. 2018; Available from: https://www.ncbi.nlm.nih.gov/pubmed/29617521