Developing prediction models for short-term mortality after surgery for colorectal cancer using a Danish national quality assurance database

Karoline Bendix Bräuner (karob@regionsjaelland.dk)
Zealand University Hospital
Andreas W. Rosen
Zealand University Hospital
Adamantia Tsouchnika
Zealand University Hospital
Julie S. Walbech
Zealand University Hospital
Mikail Gögenur
Zealand University Hospital
Viviane A. Lin
Zealand University Hospital
Johan S. R. Clausen
Zealand University Hospital
Ismail Gögenur
Zealand University Hospital

Research Article

Keywords: colorectal cancer, machine learning, prediction model, mortality, postoperative

Posted Date: May 16th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1629192/v1

Abstract

Purpose

The majority of colorectal cancer surgeries are performed electively, and treatment is often decided at the multidisciplinary team conference. Although the average 30-day mortality rate is low, there is substantial population heterogeneity from young, healthy patients to frail, elderly patients. The individual risk of surgery can vary widely, and tailoring treatment for colorectal cancer may lead to better outcomes. This requires risk prediction that is accurate and available prior to surgery.

Methods

Data from the Danish Colorectal Cancer Group database was transformed into the Observational Medical Outcomes Partnership Common Data Model. Models were developed to predict the risk of mortality within 30, 90, and 180 days after colorectal cancer surgery using only preoperative covariates. Several machine-learning models were trained, but due to superior performance, a Least Absolute Shrinkage and Selection Operator Logistic Regression was used for the final model. Performance was assessed with discrimination (area under the receiver operating characteristic and precision recall curve) and calibration measures (calibration-in-the-large, intercept, slope, and Brier score).

Results

The cohort contained 65.612 patients operated for colorectal cancer in the period from 2001 to 2019 in Denmark. The Least Absolute Shrinkage and Selection Operator model showed an area under the receiver operating characteristic for 30-, 90- and 180-day mortality after colorectal cancer surgery of 0.871 (95% CI: 0.86–0.882), 0.874 (95% CI: 0.864–0.882) and 0.876 (95% CI: 0.867–0.883) and calibration-in-the-large of 1.01, 0.98 and 1.01 respectively.

Conclusion

The postoperative short-term mortality prediction model showed excellent discrimination and calibration using only preoperative predictors.

Highlights

- This study utilizes a national quality assurance register as cohort for the development of the prediction model
- This study presents a model with good discrimination and calibration (for 30-day mortality AUROC of 0.871 and Brier score of 0.04) using only preoperative predictors making it available in a multidisciplinary team treatment conference

Introduction

Colorectal cancer (CRC) is the third most common malignant neoplastic disease in the world, with an incidence of 1.8 million patients and 935.000 deaths per year [1]. The only definitive cure is surgery; however, exposure to surgery is also related to risk of adverse events related to morbidity, dependency and ultimately mortality.

The balance between the short and long term beneficial effects of surgery and potential harms is at the very core of any decision when scheduling patients for surgery. The correct identification of patients who face a higher risk of surgery-related complications can lead to a more optimized treatment plan for the individual patient, facilitating shared decision making and potentially decreasing perioperative morbidity and mortality through interventions before surgery [2]. The potential benefit of prehabilitation and optimization will most likely be in patients with limited physical resources, who generally have an increased risk of adverse outcomes after surgery [2, 3]. At the same time, in patients with a very low risk of complications, a good

prediction model that can identify patients with very low risk of mortality can lead to accelerated treatment trajectories both before and after surgery.

Although there are a number of clinical prediction models currently in existence predicting short-term mortality, very few are actually used in a clinical context, and even fewer only include preoperative information, which is a prerequisite if they are to be used before surgery e.g. in the multidisciplinary team (MDT) setting [4, 5]. We aimed to develop a prognostic clinical prediction model for short-term postoperative mortality after colorectal cancer surgery, including only preoperative predictors in order to address the unmet need of risk assessment prior to surgery.

Materials And Methods

Data sources

In Denmark, access to register data does not require ethical approval. Processing of health register data was filed under the research inventory of Region Zealand (record number: REG-047-2020). Quality assurance data from the Danish Colorectal Cancer Group's database (DCCG) were obtained from The Danish Clinical Quality Program. DCCG is a national quality assurance database for diagnosis and treatment of patients with CRC [6]. The registry contains data for all patients diagnosed with CRC, and having a contact at a surgical department in Denmark since May 1st 2001, and has coverage of > 95% of patients [7]. The database contains detailed information of demography, cancer diagnosis, surgery, oncological treatment, and patient outcomes [6] on over 76.000 patients. Data from DCCG was transformed into the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) and Danish vocabularies were translated into standard vocabularies [8, 9].

Study population

We included patients over the age of 18 from the DCCG database who had undergone CRC surgery from May 1st 2001 until December 31st 2019.

Outcomes

We assessed all-cause mortality during the times at risk (TAR), 30, 90, and 180 days after surgery.

Statistical analyses

The open-source tool ATLAS, provided by the Observational Health Data Science and Informatics (OHDSI) community [9], was used for model development. Data were randomly split into a training set used for model development, containing 75% of patients, and a test set used for internal validation containing 25% of patients. As such, models were trained using a random 75% and 25% training and test or internal validation set using 3-fold cross-validation to optimize hyperparameter settings [9]. The ATLAS version used was 2.9.0, and models were trained using R v. 4.0.0 with the 'PatientLevelPrediction' package v. 4.3.7 and Anaconda3 v. 4.4.0 with Python v. 3.6.10. We trained Least Absolute Shrinkage and Selection Operator (LASSO) Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Machine, K-Nearest Neighbor, multilayer perception Neural Network, and AdaBoost models. We assessed the performance of the best-performing model using Area Under the Receiver Operating Characteristic (AUROC) [10] and Area Under the Precision-Recall Curve (AUPRC) and calibration using calibration-in-the-large, calibration slope, calibration intercept, and Brier score [11].

We trained and tested the models, using both a simple model containing only sex and age as predictors and a more complex model containing all preoperative covariates available in DCCG to assess whether or not adding more clinical granularity improved performance. Missing data in the source database was considered to mean that the patient did not have a record of this (for instance a given diagnosis), which means that there is a risk of misclassification of patients with missing data, however, age and sex are mandatory fields in OMOP-CDM and will therefore never be missing.

The study reporting adheres to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines [12].

Results

Participants

A total of 65,612 patients underwent surgery for colorectal cancer from 2001-2019, consisting of 53.4 % female patients with a median age of 71 years. Most patients (86.9 %) underwent elective surgery. The incidence of 30-day mortality was 5.42%, 90-day mortality was 8.53%, and 180-day mortality was 11.42%. Patient characteristics can be viewed in table 1.

Model development

Out of the models tested, the LASSO Logistic Regression performed best in the three times at risk with excellent discrimination (figure 1 and 2a-c) and good calibration (figure 3a-c)

Model performance

Discrimination of the LASSO regression model is presented as AUROC and AUPRC in table 4 and figures 2a-c, and calibration is shown in table 4 and figures 3a-c, and top 15 positive and negative weighted covariates from the 30-day mortality model are seen in tables 2 and 3. The 90- and 180-day models had similar covariate distribution, and the tables can be found as supplementary tables 2 and 3. For all three LASSO mortality models, we found excellent discrimination with AUROC og 0.871, 0.874, and 0.876 respectively. The AUPRC of 0.35, 0.44, and 0.54 respectively should be considered good as a AUPRC should be larger than the incidence of the outcome, and AUPRC here are 7-, 5-, and 5-fold higher (table 1). Calibration was also excellent seen by calibration-in-the-large from 0.98-1.02, calibration slope very close to 1 and calibration intercept and Brier score near 0 for all times at risk seen in table 4 and through assessment of weak calibration plots in figures 5-7.

When comparing the complex models with the models using only age and sex as covariates, we saw markedly better performance in the complex models, which was excellent in terms of discrimination (AUROC > 0.8) and good in terms of calibration, whereas the simple model only showed moderate to fair discrimination (AUROC > 0.6 and > 0.7) and markedly lower AUPRC.. Calibration measures were more similar than discrimination, as seen in table 4.

Table 4. *Cohort, calibration, and discrimination metrics for 30-, 90- and 180-day post-operative mortality using LASSO Logistic Regression*

Discussion And Conclusion

We trained prediction models for short-term mortality after colorectal cancer surgery based solely on preoperative covariates with excellent discrimination and good calibration. Discrimination in terms of AUROC ranged from 0.871–0.876 and AUPRC from 0.35–0.54 and calibration ranging from calibration-in-the-large 0.98–1.01, calibration slope 1.001–1.02, calibration intercept − 0.06 − 0.05, Brier score 0.04–0.07, and with solid calibration plots as seen in Figs. 5–7. Compared to models based on only age and sex as predictors, the data-driven prediction models showed vastly better performance. Based on the calibration plots, the model slightly underpredicted risks for patients with more than 50% risk of mortality.

All predictors used in the prediction model could be available at a preoperative MDT-conference. The risk factors for short-term postoperative mortality are aligned with the current literature, namely that increasedage, high American Society of Anesthesiology (ASA) score, exploratory procedures, and poor tumor differentiation were risk factors for mortality [13]. We found that predictors such as young age, low World Health Organization Performance Status (WHO PS), low ASA score, and slightly overweight body mass index (BMI) were associated with a lower risk of death during the time at risk.

Designing prediction models targeted for clinical use is not a new phenomenon. The most well-known surgical risk assessment tool is the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) Surgical Risk Calculator. Discriminative accuracy for 30-day mortality showed an AUROC of 0.944 and a Brier score of 0.011 [14], however in

validation for colorectal cancer patients performance was somewhat lower with AUROC of 0.86 and a Brier score of 0.018. Comparably Fazio et al. designed a 30-day mortality model after colorectal surgery with an AUROC of 0.801 [15], van der Sluis et al. created the Identification of Risk in Colorectal Surgery score with an AUROC of 0.83 [16], and van den Bosch et al. designed a 30-day mortality model with an AUROC of 0.82 [17]. Generally, most studies do not include as many performance metrics, and a majority show no calibration measures such as calibration-in-the-large, intercept, and calibration slope. Brier score has previously been critizised for not being an optimal measure of performance and calibration in clinical models [18], and other parameters such as calibration-in-the-large are considered essential for external validation [19]. The four studies above all defined the covariates of the model initially, however, our approach to the model was to provide the LASSO model with all available covariates and let it exclude all irrelevant covariates in order to use all covariates, that affected the prediction. This model requires 50 variables in the shape of 114–142 covariates, which are either positive or negative, and this is a large amount of variables to input into the model. However, this issue could be addressed through automation of data retrieval for the models through software interfaces to the electronic health record (EHR). Although having a large number of covariates might be impractical for input purposes, including a large number of variables with a data-driven approach minimizes bias from experts including variables assumed to be important without considering all possible options. On the other hand, including variables without a clinician set boundary may lead to bigger variance of covariates and covariate weights. Fazio et al. had 6 preoperative covariates, van der Sluis et al. used 8 pre-, intra- and postoperative covariates, and the ACS NSQIP surgical risk calculator included 21 preoperative covariates. In comparison, we found 50 weighted preoperative predictors.

We view the use of our model as a tool to estimate mortality risk and tailor different patient treatment trajectories. This is because the current treatment guidelines for colorectal cancer leads some patients to overtreatment and some to undertreatment – both with unnecessarily high risk for the patient. The model should be viewed as a decision-support tool rather than a decision-making tool, where the individual patient risks should be put into context by experienced clinicians and fuel multidisciplinary treatment approaches.

Knowledge about individual risks of mortality shortly after surgery can support the MDT-conferences in making individualized treatment plans, which takes all relevant risk factors into consideration. This personalization of treatment to risk profiles may limit both over- and undertreatment and consequences thereof.

A significant limitation of this study is the lack of external validation, which is essential for testing model generalizability and has been shown to improve clinicians' trust in the model and its predictions [20]. Also, due to the complexity of the treatment of colorectal cancer and the multitude of different variables in DCCG, some variables may be proxy for outcomes or actions in the patient course, which can lead to multicollinearity [21]. However, this is partly addressed using LASSO Logistic Regression, which considers whether or not multicolleniarity seems to occur between variables and downscales their predictive weight [22].

Strengths of this study include the development of a prediction model based on a large national validated quality assurance health database including more than 95% of all patients with colorectal cancer in Denmark, and that the model only includes preoperative data, making the model available as a clinical decision support tool in an the preoperative setting. The utilization of OMOP-CDM allows for future external validation and enrichment of data from other databases.

In conclusion, we found that designing a short-term postoperative mortality model for outcomes after colorectal surgery using a data-driven approach and utilizing only preoperative covariates is feasible and leads to models with excellent discrimination and good calibration.

Abbreviations

American Society of Anesthesiology (ASA)

Area under precision recall curve (AUPRC)

Area under the receiver operating characteristic (AUROC)

Body mass index (BMI)

Common data model (CDM)

Colorectal cancer (CRC)

Danish Colorectal Cancer Group database (DCCG)

Least Absolute Shrinkage and Selection Operator (LASSO)

Multidisciplinary team (MDT)

Negative predictive value (NPV)

Observational Health Data Science and Informatics (OHDSI)

Observational Medical Outcomes Partnership (OMOP)

Positive predictive value (PPV)

Systemized Nomenclature of Medicine (SNOMED)

Time at risk (TAR)

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)

World Health Organization performance status (WHO PS)

Declarations

Competing interests and funding

All authors have approved the submission of this paper. All authors declare no conflicts of interest. No authors have received financial grants or have financial relationships.

This paper in an earlier format has been presented as a poster at the Observational Health Data Science and Informatics Symposium 2021 (online convention) running from Sunday September 12th 2021 to Wednesday September 15th 2021 and as a lightning talk on the Danish Surgical Society annual meeting in Copenhagen, Denmark running from Thursday November 18th 2021 to Friday November 19th 2021

Acknowledgements

The authors thank The Danish Clinical Quality Program and Danish Colorectal Cancer Group for access to their outstanding data. We also sincerely thank the European Health Data Evidence Network (EHDEN), edenceHealth, and Computerome for contributions and support during the project. Lastly, we thank Peter Rijnbeek and Iannis Drakos for sparring and advice in the process.

Author contributions

All authors contributed to the study design. Karoline Bendix Bräuner, Julie Sparholt Walbech, and Adamantia Tsouchnika did the statistical analyses. Karoline Bendix Bräuner, Andreas Weinberger Rosen, Viviane Annabell Lin, Mikail Gögenur, and Johan Stub Rønø Clausen defined the research questions and interpreted the results. Karoline Bendix Bräuner and Julie Sparholt Walbech prepared tables and figures. Karoline Bendix Bräuner wrote the main manuscript text. All authors have reviewed the manuscript and given approval for submission to International Journal of Colorectal Disease.

Conflicts of interest

All authors have approved the submission of this paper. All authors declare no conflicts of interest.

References

1. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. Dis Colon Rectum 2010;53:1099. https://doi.org/10.1007/DCR.0b013e3181d60a51.
2. Bojesen RD, Degett TH, Dalton SO, Gögenur I. High World Heath Organization Performance Status Is Associated With Short And Long-term Outcomes After Colorectal Cancer Surgery. Dis Colon Rectum 2021:58–66. https://doi.org/10.1097/DCR.0000000000001982.
3. Bojesen RD, Jørgensen LB, Grube C, Skou ST, Johansen C, Dalton SO, et al. Fit for Surgery — feasibility of short-course multimodal individualized prehabilitation in high-risk frail colon cancer patients prior to surgery. Pilot Feasibility Stud 2022:1–13. https://doi.org/10.1186/s40814-022-00967-8.
4. Palmer G, Martling A, Cedermark B, Holm T. Preoperative tumour staging with multidisciplinary team assessment improves the outcome in locally advanced primary rectal cancer. Colorectal Dis 2011:1361–9. https://doi.org/10.1111/j.1463-1318.2010.02460.x.
5. Rosander E, Holm T, Sjövall A, Hjern F, Weibull CE, Nordenvall C. Preoperative multidisciplinary team assessment is associated with improved survival in patients with locally advanced colon cancer; a nationwide cohort study in 3157 patients. Eur J Surg Oncol 2021;47:2398–404. https://doi.org/10.1016/j.ejso.2021.05.008.
6. Ingeholm P, Gögenur I, Iversen LH. Danish colorectal cancer group database. Clin Epidemiol 2016;8:465–8. https://doi.org/10.2147/CLEP.S99481.
7. Klein MF, Gögenur I, Ingeholm P, Njor SH, Iversen LH, Emmertsen KJ. Validation of the Danish Colorectal Cancer Group (DCCG.dk) database - on behalf of the Danish Colorectal Cancer Group. Color Dis Off J Assoc Coloproctology Gt Britain Irel 2020;22:2057–67. https://doi.org/10.1111/codi.15352.
8. Vogelsang RP, Bojesen RD, Hoelmich ER, Orhan A, Buzquurz F, Cai L, et al. Prediction of 90-day mortality after surgery for colorectal cancer using standardized nationwide quality-assurance data. BJS Open 2021;5. https://doi.org/10.1093/bjsopen/zrab023.
9. Observational Health Data Sciences and Informatics. The Book of OHDSI 2019:1–470.
10. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010;5:1315–6. https://doi.org/10.1097/JTO.0b013e3181ec173d.
11. Rufibach K. Use of Brier score to assess binary predictions. J Clin Epidemiol 2010;63:938–9. https://doi.org/10.1016/j.jclinepi.2009.11.009.
12. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med 2015;162:55–63. https://doi.org/10.7326/M14-0697.
13. Iversen LH, Ingeholm P, Gögenur I, Laurberg S. Major reduction in 30-day mortality after elective colorectal cancer surgery: A nationwide population-based study in Denmark 2001–2011. Ann Surg Oncol 2014;21:2267–73. https://doi.org/10.1245/s10434-014-3596-7.
14. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, et al. Surgical Risk Calculator: A Decision Aide and Informed Consent Tool for Patients and Surgeons. J Am Coll Surg 2013;217(5):833–842.e3. https://doi.org/10.1016/j.jamcollsurg.2013.07.385.Development.
15. Fazio VW, Tekkis PP, Remzi F, Lavery IC. Assessment of operative risk in colorectal cancer surgery: The cleveland clinic foundation colorectal cancer model. Dis Colon Rectum 2004;47:2015–24. https://doi.org/10.1007/s10350-004-0704-y.
16. van der Sluis FJ, Espin E, Vallribera F, de Bock GH, Hoekstra HJ, van Leeuwen BL, et al. Predicting postoperative mortality after colorectal surgery: A novel clinical model. Color Dis 2014;16:631–9. https://doi.org/10.1111/codi.12580.

17. Van Den Bosch T, Warps ALK, De Nerée Tot Babberich MPM, Stamm C, Geerts BF, Vermeulen L, et al. Predictors of 30-Day Mortality among Dutch Patients Undergoing Colorectal Cancer Surgery, 2011–2016. JAMA Netw Open 2021;4:2011–6. https://doi.org/10.1001/jamanetworkopen.2021.7737.
18. Assel M, Sjoberg DD, Vickers AJ. The Brier score does not evaluate the clinical utility of diagnostic tests or prediction models. Diagnostic Progn Res 2017;1:19. https://doi.org/10.1186/s41512-017-0020-3.
19. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38. https://doi.org/10.1097/EDE.0b013e3181c30fb2.
20. Scott I, Carter S, Coiera E. Clinician checklist for assessing suitability of machine learning applications in healthcare. BMJ Heal Care Informatics 2021;28. https://doi.org/10.1136/bmjhci-2020-100251.
21. Kim JH. Introduction Multicollinearity and misleading statistical results KJA. Korean J Anesth 2019:558–69. https://doi.org/10.4097/kja.19087.
22. Cepeda MS, Reps J, Ryan P. Finding factors that predict treatment-resistant depression: Results of a cohort study. Depress Anxiety 2018;35:668–73. https://doi.org/10.1002/da.22774.

Tables

Table 1. Demographic and clinical parameters for patients included in the cohort

Characteristic	Count (n = 65,629)	%
Number of events		
30-day mortality incidence	3,557	5.4
90-day mortality incidence	5,596	8.5
180-day mortality incidence	7,496	11.4
Basic information		
Male	35,073	53.4
Age	Mean (SD*) 70,54 (11,2)	Median (IQR* 25,75) 71 (64,75)
Lifestyle and comorbidity		
ASA* Score		
1	13,850	21.1
2	33,598	51.2
≥ 3	15,478	23.6
Missing	*2,703*	*4.1*
WHO* Performance status		
0	14,655	22.3
1	5,519	8.4
≥ 2	2,380	3.7
Missing	*43,075*	*65.6*
Charlson's Comorbidity Index		
0	41,567	63.3
1	10,872	16.6
2	6,514	9.9
≥ 3	6,674	10.2
Missing	*2*	*0*
Alcohol consumption per week		
0 units	11,898	18.1
1-14 units	29,503	45
15-21 units	3,874	5.9
≥ 22 units	3,602	5.5
Missing	*16,752*	*25.5*
Body Mass Index		

≤ 18.5	1,891	2.9
18.5-25	23,815	36.3
25-30	18,815	28.7
30-35	6,279	9.6
≥ 35	2,122	3.2
Missing	*12,707*	*19.3*
Smoking status		
Smoker	10,518	16
Previous smoker	21,143	32.2
Never smoked	19,298	29.4
Missing	*14,670*	*22.4*
Cancer topography		
Colon cancer	44,550	67.9
Rectum cancer	20,850	31.8
Missing	*229*	*0.3*
Tumor specific details		
T stage*		
T0	445	0.7
T1	4,629	7.1
T2	3,650	5.6
T3	8,765	13.4
T4	2,644	4
Tx	3,054	4.7
Missing	*42,442*	*64.5*
N stage*		
N0	7,272	11.1
N1	3,411	5.2
N2	2,400	3.7
Nx	3,467	5.3
Missing	*49,079*	*74.8*
M stage*		
M0	54,374	82.9
M1	10,422	15.9
Missing	833	1.2
Treatment details		
Family history of malignant neoplasm of gastrointestinal tract	16,083	24.5

Treatment intent		
Curative	48,443	73.8
Palliative	4,135	6.3
Missing	*13,051*	*19.9*
Priority of surgery		
Elective surgery	57,083	86.9
Emergency operation	7,590	11.5
Missing	*956*	*1.6*
Surgical approach		
Laparoscopy	26,587	40.5
Laparotomy	31,407	47.9
Endoscopy	3,875	5.9
Missing	*3,760*	*5.7*

**SD = Standard deviation, IQR = Interquartile range, ASA = American Society of Anesthesiology score, WHO = World Health Organization, T = Tumor category, N = Node category, M = Metastasis category, pMMR = proficient mismatch repair, dMMR = deficient mismatch repair, EMR = endoscopic mucosal resection, RFA = Radio frequency ablation*

Table 2. Top 15 covariates associated with increased risk for 30-day mortality according to the LASSO model (for full covariate table, see supplementary table 1). In total, 114, 124, and 142 preoperative covariates were used in the 30-, 90- and 180-day mortality models respectively.

Covariate name	Covariate value	Covariate count	Covariate mean	Covariate standard deviation
age group: 95 - 99	1.173	190	0.003	0.054
age group: 90 - 94	1.021	1317	0.02	0.14
age group: 85 - 89	0.884	4622	0.07	0.256
age group: 80 - 84	0.594	8373	0.128	0.334
ASA Score 4*	1.098	1300	0.02	0.139
ASA Score 3*	0.391	14118	0.215	0.411
Primary malignant neoplasm of splenic flexure of colon	0.335	1939	0.03	0.169
MX category*	0.479	8961	0.137	0.343
Emergency operation	0.484	7584	0.116	0.32
Only exploratory surgery, diagnostic laparoscopy or exploratory laparotomy	0.932	300	0.005	0.067
Endoscopic insertion of permanent colonic stent	0.488	862	0.013	0.114
Ileocolic resection	0.346	130	0.002	0.044
Defunctioning stoma	0.559	2394	0.036	0.187
Tumor perforation, open perforation	0.555	1442	0.022	0.147
Gastrointestinal perforation	0.471	945	0.014	0.119

**ASA = American Society of Anesthesiology score, MX = Metastasis category unknown*

Table 3. Top 15 covariates associated with decreased risk for 30 day mortality based on the LASSO model (for full covariate table, see supplementary table 1)

Covariate name	Covariate value	Covariate count	Covariate mean	Covariate standard deviation
age group: 50 - 54	-1.034	2914	0.044	0.206
age group: 55 - 59	-0.883	4849	0.074	0.262
age group: 60 - 64	-0.593	7508	0.114	0.318
ASA Score 1*	-0.889	13850	0.211	0.408
ASA Score 2*	-0.37	33595	0.512	0.5
WHO Performance status 0*	-0.484	14655	0.223	0.416
Body Mass Index(by algorithm) >25 <=30	-0.581	18811	0.287	0.452
Body Mass Index(by algorithm) >18.5 <=25	-0.393	23809	0.363	0.481
Never smoked tobacco	-0.433	19297	0.294	0.456
Curative - procedure intent	-0.355	48435	0.738	0.44
Endoscopic procedure	-0.939	3874	0.059	0.236
Laparoscopy	-0.352	26577	0.405	0.491
Abdominoperineal resection	-0.373	5205	0.079	0.27
Other local resection including polypresection or EMR*	-0.475	2152	0.033	0.178
Local macroradical excision of colorectal tumor	-0.568	22133	0.337	0.473

**ASA = American Society of Anesthesiology score, WHO = World Health Organization, EMR = Endoscopic mucosal resection*

Table 4. Cohort, calibration, and discrimination metrics for 30-, 90- and 180-day post-operative mortality using LASSO Logistic Regression

	Patients	Number of outcomes	Predictors	AUROC*	AUPRC*	Calibration in large	Calibration-slope	Calibration-intercept	Brier score
30-day mortality	65612	3557	50	0.871 (0.86-0.88)	0.35	1.01	1.02	0.05	0.04
30-day mortality (simple)	65612	3557	2	0.728 (0.71-0.75)	0.13	0.99	1.049	0.13	0.05
90-day mortality	65612	5596	50	0.874 (0.87-0.88)	0.44	0.98	1.001	0.023	0.06
90-day mortality (simple)	65612	5596	2	0.689 (0.68-0.70)	0.18	1.00	0.97	-0.06	0.07
180-day mortality	65612	7496	50	0.876 (0.87-0.88)	0.54	1.01	1.004	-0.006	0.07
180-day mortality (simple)	65612	7496	2	0.693 (0.68-0.71)	0.23	1.01	1.08	0.14	0.10

**Area under the receiver operating characteristic (AUROC), Area under precision recall curve (AUPRC)*

Figures

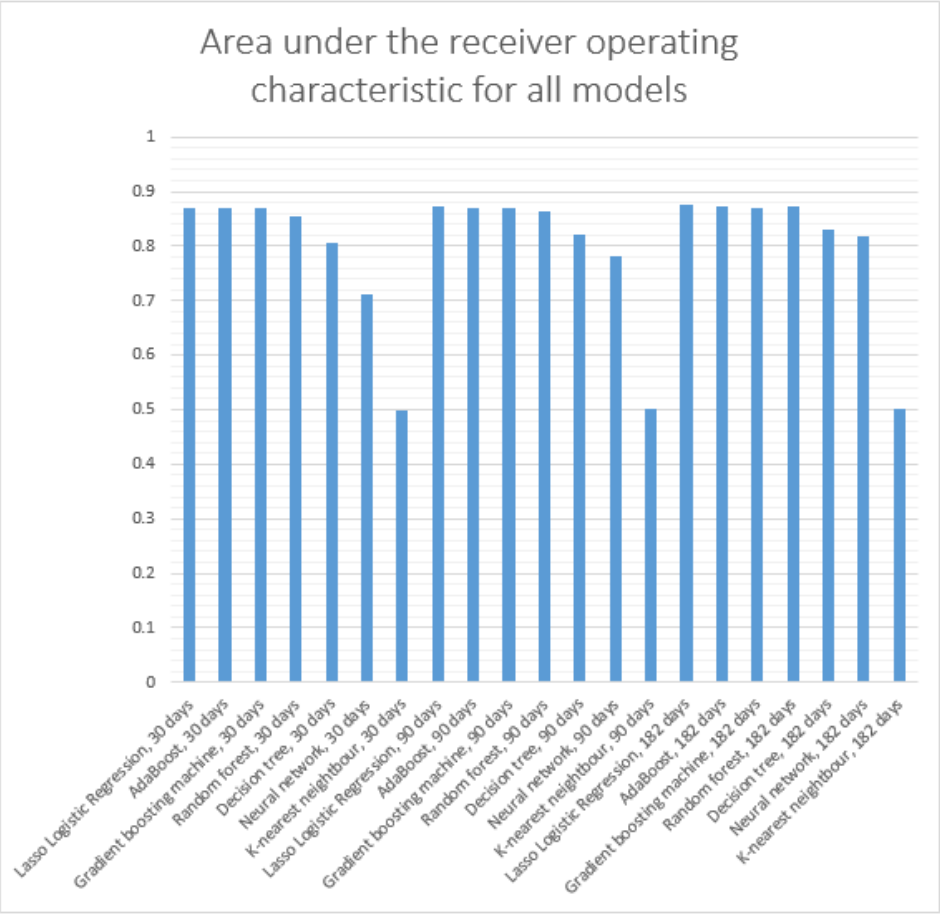

Figure 1

Area under the receiver operating characteristic (AUROC)) for each model and analysis

Figure 2

a-c. Receiver operation characteristic (ROC) curves and precision recall curves for all times at risk in LASSO Logistic Regression **2a.** Receiver operation characteristic (ROC) curve and precision recall curve for 30 day post-operative mortality using LASSO Logistic Regression **2b.** Receiver operation characteristic (ROC) curve and precision recall curve for 90 day post-operative mortality using LASSO Logistic Regression **2c.** Receiver operation characteristic (ROC) curve and precision recall curve for 180 day post-operative mortality using LASSO Logistic Regression

Figure 3

a-c. Calibration plots for all times at risk from LASSO Logistic Regression model

3a. Calibration plot for 30 day post-operative mortality using LASSO Logistic Regression for the test-set of data **3b.** Calibration plot for 90 day post-operative mortality using LASSO Logistic Regression for the test-set of data **3c.** Calibration plot for 180 day

post-operative mortality using LASSO Logistic Regression for the test-set of data

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BruneretalSUPPTABLE1.docx