



# Training prediction models for individual risk assessment of postoperative complications after surgery for colorectal cancer

V. Lin<sup>1</sup> · A. Tsouchnika<sup>1</sup> · E. Allakhverdiiev<sup>1</sup> · A. W. Rosen<sup>1</sup> · M. Gögenur<sup>1</sup> · J. S. R. Clausen<sup>1</sup> · K. B. Bräuner<sup>1</sup> · J. S. Walbech<sup>1</sup> · P. Rijnbeek<sup>2</sup> · I. Drakos<sup>1</sup> · I. Gögenur<sup>1</sup>

Received: 13 January 2022 / Accepted: 20 April 2022 / Published online: 20 May 2022  
© Springer Nature Switzerland AG 2022

## Abstract

**Background** The occurrence of postoperative complications and anastomotic leakage are major drivers of mortality in the immediate phase after colorectal cancer surgery. We trained prediction models for calculating patients' individual risk of complications based only on preoperatively available data in a multidisciplinary team setting. Knowing prior to surgery the probability of developing a complication could aid in improving informed decision-making by surgeon and patient and individualize surgical treatment trajectories.

**Methods** All patients over 18 years of age undergoing any resection for colorectal cancer between January 1, 2014 and December 31, 2019 from the nationwide Danish Colorectal Cancer Group database were included. Data from the database were converted into Observational Medical Outcomes Partnership Common Data Model maintained by the Observation Health Data Science and Informatics initiative. Multiple machine learning models were trained to predict postoperative complications of Clavien–Dindo grade  $\geq 3B$  and anastomotic leakage within 30 days after surgery.

**Results** Between 2014 and 2019, 23,907 patients underwent resection for colorectal cancer in Denmark. A Clavien–Dindo complication grade  $\geq 3B$  occurred in 2,958 patients (12.4%). Of 17,190 patients that received an anastomosis, 929 experienced anastomotic leakage (5.4%). Among the compared machine learning models, Lasso Logistic Regression performed best. The predictive model for complications had an area under the receiver operating characteristic curve (AUROC) of 0.704 (95%CI 0.683–0.724) and an AUROC of 0.690 (95%CI 0.655–0.724) for anastomotic leakage.

**Conclusions** The prediction of postoperative complications based only on preoperative variables using a national quality assurance colorectal cancer database shows promise for calculating patient's individual risk. Future work will focus on assessing the value of adding laboratory parameters and drug exposure as candidate predictors. Furthermore, we plan to assess the external validity of our proposed model.

**Keywords** Colorectal cancer · Surgery · Prediction model · Complications · Machine learning

## Introduction

Postoperative complications in colorectal cancer surgery are associated with increased morbidity and mortality as well as risk for cancer recurrence [1–3].

The ability to identify patients at high risk for complications has potential for patient optimization prior to surgery, increased surveillance postoperatively and can aid in better informed decision-making for both patient and surgeon in guiding surgical treatment. The multidisciplinary team (MDT) meeting, where patients with newly diagnosed cancer are discussed and their surgical treatment is planned, could benefit from calculated risk assessments.

Factors that are associated with an increased risk for developing severe postoperative complications or anastomotic leakage have been reported in previous studies [4–6]. However, assessing the full phenotype of a patient to quantify an individual's risk is more multifaceted. Machine learning (ML) algorithms are increasingly being utilized

✉ V. Lin  
vial@regionsjaelland.dk

<sup>1</sup> Center for Surgical Science, Department of Surgery, Zealand University Hospital Køge, Lykkebækvej 1, 4600 Køge, Denmark

<sup>2</sup> Department of Medical Informatics, Erasmus University Medical Centre, Rotterdam, The Netherlands

in medical research due to their ability to capture complex relationships between a multitude of variables, including in surgical risk prediction [7].

The aim of our study was to train and internally validate models to predict the occurrence of complications, as well as anastomotic leakage specifically, after resection for colorectal cancer. Using only variables that are available prior to surgery from the Danish Colorectal Cancer Group (DCCG) database, we tested whether they can achieve sufficient discriminative power to be used in a clinical setting.

## Materials and methods

### Data sources

Prospectively collected patient data for building the prediction models were taken from the national quality assurance ‘Danish Colorectal Cancer Group’ (DCCG) database, that has recorded information on over 95% of all patients that have received a colorectal cancer diagnosis in Denmark since 2001 [8]. It consists of demographic data as well as detailed information on comorbidities, tumor stage and localization, chemotherapy, procedure type and resection, and whether intra- and or postoperative complications occurred. The DCCG has an overall data completeness of > 96%, last validated in 2020 with an accuracy of 95% [9]. While the registration of postoperative complications has been part of the DCCG registry from inception, more detailed variables were introduced in 2014, using the Clavien–Dindo classification of severity of complications [10]. As the outcome was defined as postoperative complications of Clavien–Dindo grade 3B or higher, only this subset of data was used.

Source vocabulary data from the DCCG were transformed into standard vocabularies, primarily the systematized nomenclature of medicine clinical terms (SNOMED CT) and by using custom concepts where granularity was lacking in currently available standard vocabularies. Patient sensitive information was encrypted and data were de-identified. Data were standardized to the observational medical outcomes partnership (OMOP) common data model (CDM) and the standardized vocabularies [11, 12]. Data conversion, transformation and quality control (QC) were achieved using the observational health data sciences and informatics (OHDSI) open source tools guided by the European health data evidence network (EHDEN) and in collaboration with EdenceHealth NV (Veldkant 33A, Kontich, Belgium).

The study was approved by the Region Zealand (REG-047-2020). The study results are presented in accordance with the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement [13].

### Patient population

All patients over 18 years of age undergoing any resection for colorectal cancer between January 1, 2014 and December 31, 2019 from the nationwide DCCG database were included. Seventy-five percent of the patients were randomly selected for model training, while the remaining 25% were used for internal model validation, following a three-fold cross-validation strategy for hyper-parameter optimization. End of follow-up was 30 days after surgery. For the prediction models analyzing anastomotic leakage, only those patients receiving a gastrointestinal anastomosis were included. This included patients that received more than one resection of bowel segment, with both an intestinal anastomosis and an ostomy performed within one surgery. Table 1 shows the patient demographics and their preoperative characteristics used for training the models. In order for models to be utilized for all patients undergoing colorectal surgery, the data set was not limited to specific surgeries. Rather, when inputting patient data into the model, the model can predict outcomes for a variety of open, laparoscopic and endoscopic resections in emergent and elective settings.

While having predicted risks available prior to surgery is mostly relevant to elective surgery and the MDT setting, it was decided that models should still be able to calculate the risk for postoperative complications for emergent cases as well. Risks can still be calculated before surgery and the prediction may have an impact on surgical strategy.

### Outcomes

The primary outcome of this study was defined as a patient experiencing a postoperative complication of grade 3B or higher as defined by the Clavien–Dindo postoperative complication classification within 30 days after surgery for colorectal cancer [10]. This classification grades the severity of the complication by the necessity, manner, and invasiveness of a resulting intervention. Thus, a patient requiring a postoperative intervention under general anesthesia, admission to the intensive care unit (ICU) or death within 30 days, was considered to have experienced the target outcome.

A secondary prediction model was designed for patients receiving an anastomosis during surgery, and their risk of experiencing anastomotic leakage within 30 days after surgery. While those anastomotic leaks that were the cause of a Clavien–Dindo grade 3B or higher complication were included in the previous outcome, we wondered if we could predict anastomotic leakage specifically. This could facilitate the decision of whether an anastomosis or

**Table 1** Patient cohort used for prediction model building and their preoperative characteristics

	Patients used for prediction of postoperative complications of CD 3B or higher ( <i>n</i> = 23,907)	Patients used for prediction of anastomotic leakage ( <i>n</i> = 17,190)
Characterization of preoperative variables		
Age, years (median, IQR)	71 (64–78)	71 (64–77)
Sex (male)	13,079 (54.7%)	9255 (52.9%)
BMI, kg/m <sup>2</sup>		
< 18.5	698 (2.9%)	465 (2.7%)
18.5–30	18,141 (75.9%)	13,162 (76.6%)
> 30	4328 (18.1%)	3220 (18.7%)
Missing	740 (3.1%)	343 (2.0%)
Smoking status (smoker)	4000 (16.7%)	2799 (16.3%)
ASA score		
1	5308 (22.2%)	4093 (23.8%)
2	12,659 (53.0%)	9447 (55.0%)
≥ 3	5589 (23.4%)	3528 (20.5%)
Missing	351 (1.5%)	726 (4.2%)
Performance score		
0	14,637 (61.2%)	11,240 (65.4%)
1	5510 (23.0%)	3862 (22.5%)
≥ 2	2377 (9.9%)	1362 (7.9%)
Missing	1383 (5.8%)	726 (4.2%)
Type of cancer*		
Colon	16,684 (69.8%)	13,669 (79.5%)
Rectum	7223 (30.2%)	3521 (20.5%)
TNM staging (clinical)		
Tx	3048 (12.7%)	2138 (12.4%)
T0	445 (1.9%)	206 (1.2%)
T1	3528 (14.8%)	1931 (11.2%)
T2	3646 (15.3%)	2911 (16.9%)
T3	7022 (29.4%)	5186 (30.2%)
T4	2132 (8.9%)	1230 (7.2%)
Missing	4086 (17.1%)	3588 (20.9%)
Nx	3460 (15.2%)	2249 (13.1%)
N0	7264 (30.4%)	5316 (30.9%)
N1	3407 (14.3%)	2616 (15.2%)
N2	2396 (10.0%)	1750 (10.2%)
Missing	7380 (30.9%)	5259 (30.6%)
M0	21,219 (88.8%)	15,674 (91.2%)
M1	2515 (10.5%)	1474 (8.6%)
Missing	173 (0.7%)	42 (0.2%)
Primary procedure		
Right-sided hemicolectomy	7574 (31.7%)	7275 (42.3%)
Colectomy	546 (2.3%)	157 (0.9%)
Excision rectum	5506 (23.0%)	3822 (22.2%)
Local resection	1778 (7.4%)	0 (0.0%)
Other excision	1842 (7.7%)	209 (1.2%)
Surgical approach		
Open	4,431 (18.5%)	2,956 (17.2%)
Laparoscopic	13,951 (58.4%)	11,614 (68.3%)
Endoscopic	1,778 (7.4%)	0 (0%)

**Table 1** (continued)

	Patients used for prediction of postoperative complications of CD 3B or higher ( <i>n</i> = 23,907)	Patients used for prediction of anastomotic leakage ( <i>n</i> = 17,190)
Missing	3,747 (15.7%)	2,620 (15.2%)
Surgical intent		
Curative	22,117 (92.5%)	16,452 (95.7%)
Palliative	527 (2.2%)	250 (1.5%)
Missing	1263 (5.3%)	488 (2.8%)
Surgical priority		
Elective	21,794 (91.2%)	16,200 (94.2%)
Emergency	1990 (8.3%)	980 (5.7%)
Missing	123 (0.5%)	10 (<0.1%)
Anastomosis constructed	16,689 (69.8%)	17,190 (100%)

CD Clavien–Dindo, BMI Body mass index, ASA American society of anesthesiologists

\*Type of cancer is higher than the total number of patients in the cohort, as there were cases of synchronous tumors in both colon and rectum. Right-sided hemicolectomy includes transverse colectomy and extended right-sided colectomy. Left-sided hemicolectomy includes sigmoid colectomy

a stoma is more appropriate for the individual patient at the time of resection. Anastomotic leakage was defined as a either type A, B or C leakage diagnosed clinically, radiologically, endoscopically or surgically [14].

Prediction models compute the probability between 0 and 1 of a patient developing the predicted outcomes.

## Predictors

All variables available prior to surgery as well as decisions made prior to surgery that would be available in an MDT meeting, such as planned primary procedure, surgical approach, priority and intent, were used as candidate predictors in the prediction models. Predictors included demographic variables, tumor staging and localization, biopsy results, imaging results, known comorbidities, planned primary procedure and neo-adjuvant chemotherapy. Some predictors could be available only for certain patients, such as tumor distance from the anal verge measured by flexible or rigid endoscopy or MRI staging of lymph nodes, which would therefore only be available for risk calculations in patients with rectal cancer.

## Missing data

Missing data were considered missing at random. Standard practices in ML were followed by using one-hot encoding for categorical variables. Missing data were classified as the absence of these categorical variables.

## Statistical analysis

All preoperative variables were used as predictors in initial analysis, after which a manual feature selection was performed and variables were selected that were relevant and available at an MDT meeting. All categorical predictors were converted into binary predictors. Minimum cohort sample size calculations were performed using the ‘pmsampsize’ package v.1.1.2 in R.

For prediction model training, the ATLAS version 2.9.0 interface for OHDSI tools was used to design the study and run characterizations. The prediction models were trained using the ‘Patient Level Prediction’ package, R version 4.1.0 and Python version 3.9.6 [15, 16]. Prediction model algorithms included least absolute shrinkage and selection operator (LASSO) logistic regression, gradient boosting machines, adaboost, random forest, K nearest neighbor, multilayer perceptrons and decision trees, and models were trained using the default hyper-parameters settings. The models with the best performance on the internal validation sets were then selected based on their capability of discrimination using the AUROC and Precision-Recall curve.

Sensitivity and specificity were obtained for their respective thresholds. As there is a tradeoff between sensitivity and specificity, thresholds can be set depending on which is favored in a specific clinical setting. A threshold set for high specificity would thereby be preferred in a setting where knowing whether an individual will not get a complication, is prioritized. When requiring high diagnostic accuracy on the probability of getting an outcome such as anastomotic leakage, a high sensitivity might be preferred and a threshold

can be set accordingly, in order to identify those that might benefit from a stoma.

## Results

A total of 23,907 patients underwent surgery for colorectal cancer in Denmark between 2014 and 2019. Of these, 2,958 patients (12.4%) experienced a complication of Clavien–Dindo grade 3B or higher within 30 days after surgery. Of 17,190 patients that received a gastrointestinal anastomosis in the same timeframe, a total of 929 patients (5.4%) experienced anastomotic leakage. Table 1 shows the patient data used for prediction modeling. Minimum sample size calculation results can be found in the supplementary Tables 1 and 2.

### Prediction models

#### Postoperative complications Clavien–Dindo grade 3B or higher

The prediction model with the best discrimination and calibration for predicting postoperative complications of Clavien–Dindo grade 3B or higher used 111 preoperative predictors in a LASSO Logistic Regression model with an AUROC = 0.704 (95%CI 0.683–0.724, Fig. 1). (Table 3 in the supplement shows performance metrics for all other ML models trained). After reverting binary variables to categorical variables, the variable selection process singled out a total of 30 variables for predicting the risk of postoperative complications of Clavien–Dindo grade 3B or higher (Table 2

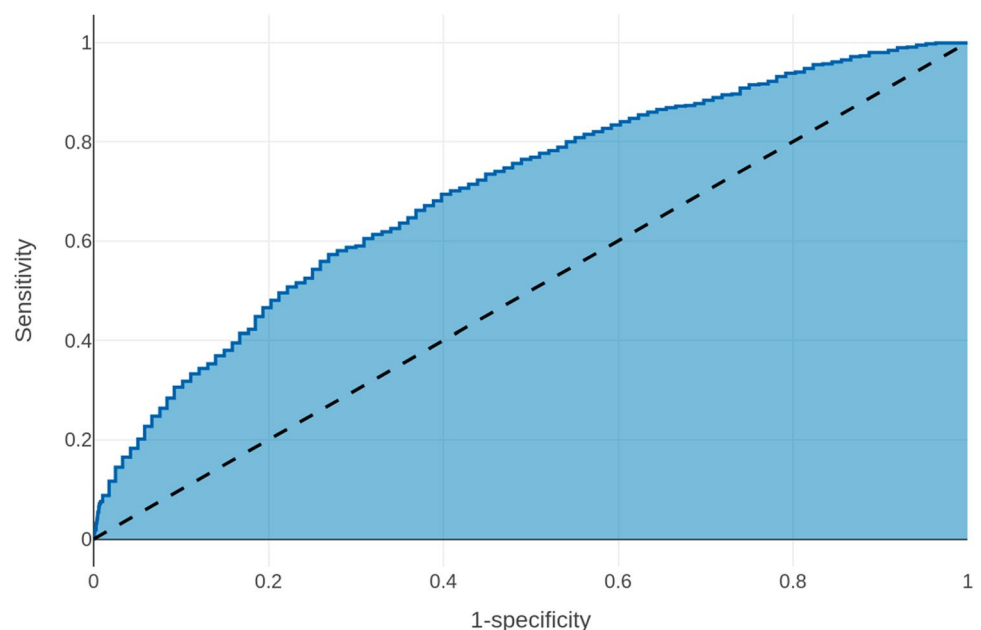
shows demographic predictors used in the model. The full list can be found in the supplementary Table 4). The area under the precision-recall curve (AUPRC) was at 0.285 (95%CI 0.252–0.317). Calibration of the validation set was good, with calibration-in-the-large 1.01 (ratio of mean predicted risk/mean observed risk), calibration-intercept 0.10 and calibration-slope 1.06. (Fig. 2). The Brier score was 0.10.

Based on regression coefficients, American Society of Anesthesiologists (ASA) scores of 3 or higher, performance status of 2 or higher, abdominal colectomies, open surgery, emergency surgery and tumor perforation had high impact on the model and were associated with a higher risk for postoperative complications. Predictors associated with a lower risk of postoperative complications were endoscopic resections, female patients, right-sided and sigmoid colectomies, ASA score of 1, curative resections, resections without creation of an ostomy, and patients with slightly elevated body mass index (BMI).

#### Anastomotic leakage

The prediction model with the best discrimination and calibration for predicting anastomotic leakage used 58 preoperative predictors in a LASSO Logistic Regression model with an AUROC = 0.690 (95%CI 0.655–0.724, Fig. 3). (Supplementary Table 5 shows performance metrics for all other ML models trained). When reverting binary variables to categorical variables, the model used a total of 27 variables for predicting the risk of anastomotic leakage (Table 3). The AUPRC was at 0.119 (95%CI 0.079–0.162). Calibration of the testing set was also good with a calibration-in-the-large

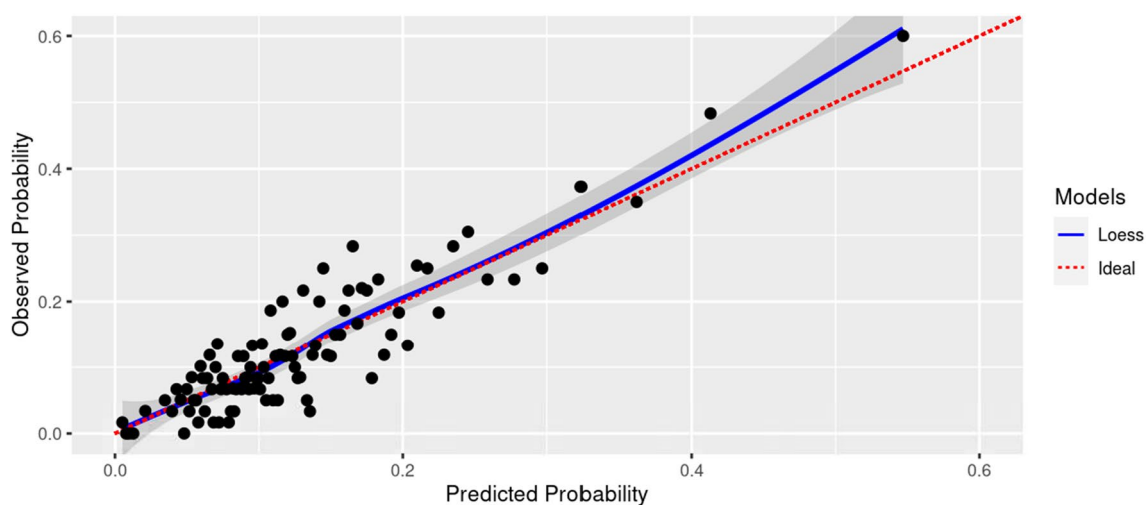
**Fig. 1** Receiver operating characteristic plot of LASSO logistic regression model for predicting postoperative complications CD 3B or higher for patients operated on between 2014–2019 CD-3B:Clavien–Dindo grade 3B



**Table 2** Demographic preoperative variables list used in the model for predicting Clavien–Dindo complications grade 3B or higher, with regression coefficients for binary predictors (Full list of predictors can be found in the supplementary table 1)

Categorical variable	Predictors contributing to lower risk	Regression coefficient	Predictors contributing to higher risk	Regression coefficient
Sex	Female	− 0.3429		
Age group (years)	40–44	− 0.1567	35–39	0.3653
	60–64	− 0.0313	45–49	0.1586
	65–69	− 0.0038	50–54	0.0894
	75–79	− 0.0004	80–84	0.1017
			85–89	0.0961
BMI, kg/m <sup>2</sup>	18.5–25	− 0.2255	< 18.5	0.0450
	25–30	− 0.2996		
	30–35	− 0.0275		
ASA Score	1	− 0.4381	3	0.1140
	2	− 0.1357	4	0.6349
Alcohol consumption (units/week)	0	− 0.0576	15–21	0.0014
	1–14	− 0.1136	> 21	0.1399
Charlson score	0	− 0.0932		
	1	− 0.0320		
	2	− 0.0837		
	3+	− 0.0187		
Performance status			0	0.0294
			1	0.1706
			2	0.3957
			3	0.5166
			4	0.2324
Smoking status	Never smoked	− 0.0932	Currently smoker	0.2884
			Ex-smoker	0.1018
Family history	Family history of colorectal cancer	− 0.1645		

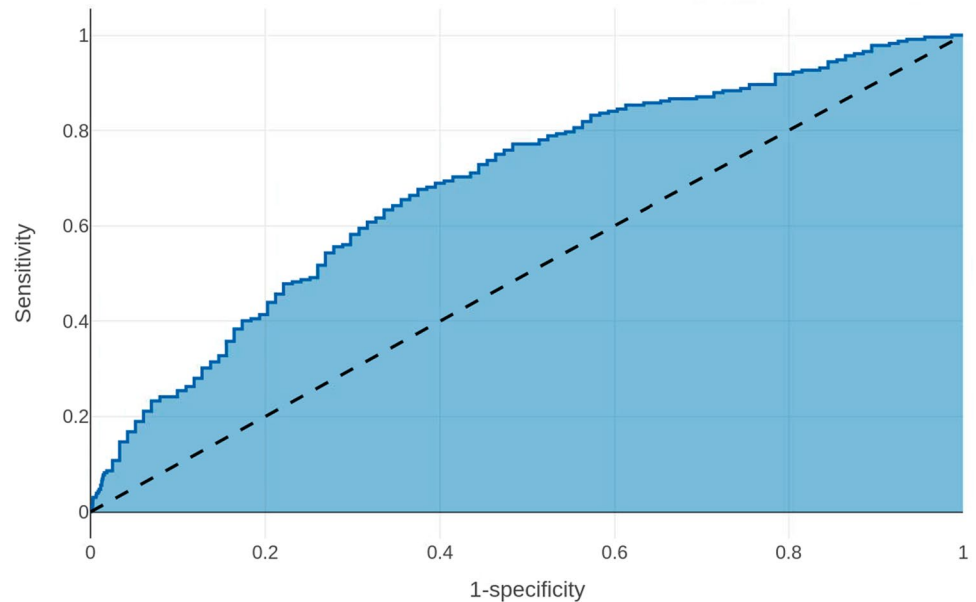
BMI Body mass index, ASA American society of anesthesiologists



**Fig. 2** Calibration plot of LASSO logistic regression model using loess algorithm for internal validation set for predicting postoperative complications CD 3B or higher for patients operated on between

2014 and 2019. Dots represent subgroups of patients and their predicted vs. observed probability of the outcome (*Loess* Locally estimated scatterplot smoothing) *CD-3B* Clavien–Dindo grade 3B

**Fig. 3** Receiver operating characteristic plot of LASSO Logistic Regression model for predicting anastomotic leakage for patients operated on and receiving a gastrointestinal anastomosis between 2014 and 2019



of 1.00, a calibration-intercept of -0.15 and calibration-slope 0.94 (Fig. 4). The Brier score was 0.05.

Patients with rectum resections were predicted to be at higher risk for anastomotic leakage, as were patients with a higher BMI, ASA score of 3, ileorectal anastomosis and smokers. Variables used as predictors for not developing anastomotic leakage were among others right-sided hemicolectomies, female patients, non-smokers, patients receiving a permanent ostomy and ASA 1 score.

The full covariate list with the intercept is shown in the supplementary Table 6.

### Thresholds

Positive and negative predictive values as well as sensitivity and specificity were calculated according to the various risk thresholds (see supplementary Tables 7 and 8).

### Discussion

We trained two prediction models using a national quality assurance database to predict the risk of postoperative complications and anastomotic leakage after colorectal cancer surgery. Only predictors that were available during the MDT meeting were included in order to enable informed decision-making prior to surgery.

Model performances using only preoperative variables available from the registry were not up to a sufficient standard for use in a clinical setting yet with an AUROC of 0.704 (95%CI 0.683–0.724) and 0.689 (95%CI 0.654–0.724) respectively for discriminative power. Interestingly, the automatically selected predictors for the postoperative

complications risk models agreed with the current literature. Variables with the highest absolute weights associated with a higher risk and used by the prediction model were ASA score 3 or higher, performance status  $\geq 2$ , emergency surgery, undergoing total colectomy and open surgery. Similarly, variables in the model identified as being associated with a lower risk for complications were endoscopic resections, ASA 1 score, female patients, and patients with a slightly elevated BMI.

The model for prediction of anastomotic leakage equally identified and utilized predictors that are known to be associated with anastomotic leakage [17]. Predictors associated with the occurrence of anastomotic leakage were patients that had rectal cancer, were smokers and with a BMI  $> 30$  kg/m<sup>2</sup>. Female patients, non-smokers, ASA 1 score and right-sided resections were predictors associated with not getting an anastomotic leak.

Besides selecting predictors that are known to be associated with postoperative complications or anastomotic leakage, our models selected additional parameters using a total of 30 and 27 categorical predictors respectively. This illustrates the complexity of individual risk assessment and how a data-driven approach can add value to predictions.

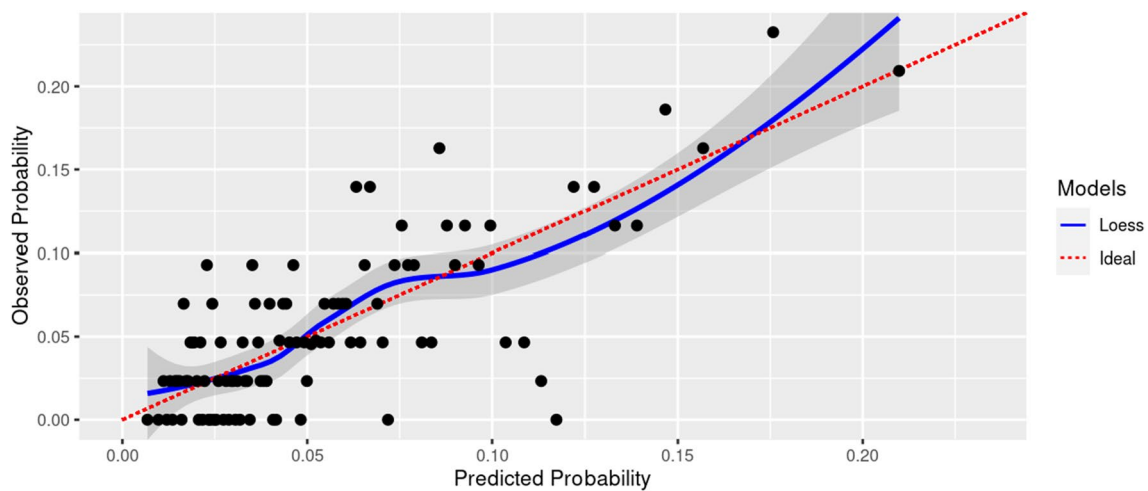
The American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) has similarly been used to construct a universal risk calculator, that can predict complications after a wide variety of surgeries. A specific machine learning model for hepatopancreatic and colorectal surgeries was also created [7, 18]. While predictive performance for the latter was very high (AUROC = 0.816) in the universal risk calculator for postoperative complications, discriminatory power varied greatly during subsequent external validation by other groups

**Table 3** Full preoperative variables list used in the model for predicting anastomotic leakage, with regression coefficients for binary predictors

Categorical variable	Predictor contributing to lower risk	Regression coefficient	Predictor contributing to higher risk	Regression coefficient
Sex	Female	− 0.5430		
Age group (years)	65–69	− 0.0048	55–59	0.116
	75–79	− 0.1074		
	80–84	− 0.0486		
	85–89	− 0.0905		
BMI kg/m <sup>2</sup>	< 18.5	− 0.0211	30–35	0.2877
	18.5–25	− 0.0065	> 35	0.4498
	25–30	− 0.0055		
ASA score	1	− 0.1347	3	0.2524
Alcohol consumption (units/week)	0	− 0.0009	15–21	0.0508
	1–14	− 0.0310		
Charlson score	2	− 0.1096	1	0.0181
Performance status			1	0.0189
Smoking status	Never smoker	− 0.2256	Current smoker	0.1352
Screening			Patient who participated in colorectal cancer screening	0.0470
Histology of biopsy	Malignant adenomatous neoplasm	− 0.0344		
Histological differentiation	Intermediate to high differentiation	− 0.0552		
MMR status	Proficient MMR status	− 0.0627	Deficient MMR status	0.2455
Preoperative endoscopic resection	Preoperative endoscopic resection of malignant tumor	− 0.1302		
Distance from tumor to anal verge	By endoscopy > 10 cm	− 0.1062		
Localization of primary tumor	Ascending colon	− 0.0256	Transverse colon	0.0091
	Hepatic flexure	− 0.1501	Sigmoid colon	0.0682
	Synchronous tumor	− 0.0260		
Localization of metastasis	Metastasis of lung	− 0.1488		
T staging	T0 stage	− 0.0432	T4 stage	0.1204
	T3 stage	− 0.0183		
	N0 stage	− 0.0808	Nx stage	0.0717
N staging	N0 stage	− 0.0808	Nx stage	0.0717
M staging	M1 stage	− 0.0213	Mx stage	0.1354
MRI T stage			MRI T3 stage	0.2275
Neoadjuvant radio- and/or chemotherapy	Neoadjuvant radio- and/or chemotherapy	− 0.0680		
Treatment intent	Compromised resection	− 0.2721	Curative procedure intent	0.0933
Surgical approach			Open surgery (laparotomy)	0.0196
Primary procedure	Transanal resection of rectum tumor	− 0.1033	Total abdominal colectomy with ileoproctostomy	0.6108
	Right colectomy	− 0.8124	Additional excision besides colon or rectum necessary	0.0592
	Extended right hemicolectomy	− 0.3896	Resection of rectum	0.8212
	Local resection of colorectal tumor	− 0.0678		
	Other colectomy	− 0.0093		
	Intraoperative treatment of metastasis	Intraoperative treatment of metastasis	− 0.1823	
Additional stoma created	Permanent stoma	− 1.1695		
	Temporary stoma	− 0.0430		

*BMI* Body mass index, *ASA* American society of anesthesiologists, *MRI* Magnetic resonance imaging, *MMR* Mismatch repair





**Fig. 4** Calibration plot of LASSO Logistic Regression model using Loess algorithm for internal validation set for predicting anastomotic leakage for patients operated on and receiving a gastrointestinal anas-

tomosis between 2014 and 2019. Dots represent subgroups of patients and their predicted vs. observed probability of the outcome (*Loess* Locally estimated scatterplot smoothing)

[19–22]. Performing an external validation of existing models such as the ACS NSQIP surgical risk calculator using our data might allow for better comparison of predictive performance between models, but is not currently possible due to some data, such as steroid use that is required for the risk calculator, not being available in our dataset.

Models to predict anastomotic leakage have previously been trained using literature reviews or expert opinion of prognostic factors as well as existing datasets [17, 23, 24]. Models that have been externally validated have used both preoperative and intraoperative variables to predict the risk for anastomotic leakage [17, 25]. Among the variables used to predict anastomotic leakage were obesity, sex, ASA score, preoperative serum total proteins, ongoing anticoagulant treatment, the occurrence of intraoperative complications, blood loss or transfusion as well as duration of operation. External validation performed relatively well, with AUROC's reported between 0.623 and 0.96 [26–29]. Using intraoperative variables for prediction of anastomotic leakage means a model is unavailable prior to surgery, but of course, it still allows the surgeon to make intraoperative treatment decisions such as creating a stoma instead of an anastomosis.

It is important to acknowledge that prediction models must always be taken into context, and predictions can, but need not force a change in treatment. The identification of a patient at high risk in itself can influence an outcome. An increased surveillance in order to identify complications early in the postoperative phase has been shown to lower mortality in high-risk patients [30]. Another option can be a form of pre-habilitation for optimizing patients' health status prior to surgery [31, 32]. Alterations in treatment, such as creating an ostomy instead of a primary anastomosis, should

always be a decision made by the surgeon and the patient. Considering the high negative predictive value of our model for predicting the risk of anastomotic leakage, having a low risk for anastomotic leakage could encourage the creation of an anastomosis. (Supplementary Table 6).

In prediction modeling, the AUROC curve is essentially a visualization of predictive discrimination across all thresholds. Defining a threshold is not a necessity for predictive modeling, as the calculation of the individual risk can be sufficient for informed decision-making. Furthermore, multiple thresholds can be set based on different clinical scenarios and patients. In the clinical setting, it is the treating colorectal surgeon with the patient that effectively sets an individual threshold for a patient. However, if a binary 'high-risk' or 'low-risk' classification is desired, thresholds can be set which can be based on a harm-benefit analysis of morbidity and mortality, health care costs as well as patient and surgeon preferences.

Strengths of this study are that data-driven predictions are based on a large, validated, well-curated, nationwide quality assurance registry with high completeness. Furthermore, models only used preoperative predictors, making them available for aiding in clinical decision-making prior to surgery at the MDT meeting. Limitations include the restriction of complications to 30 days after surgery, potentially missing out on some surgery-related complications (e.g., late anastomotic leakage), and the complexity of the prediction models based on 30 and 27 categorical preoperative predictors respectively. Use of a model in a clinical setting could require simplification with additional sacrifice of performance of the model, or an application that can access the required variables directly from electronic health records to eliminate manual entering of a large amount of

data to increase usability. Due to lack of a gold standard, we deem a clinical validation of prediction models necessary before considering them in a clinical setting. Furthermore, accuracy could also be reduced because we aimed to predict complications for a wide variety of colorectal cancer surgeries in two ‘universal’ models, versus if we had constructed separate models for colon or rectum cancer surgery, or even separate primary procedures. Training procedure-specific models could be another way forward to increase predictive performance.

Utilizing additional data sources containing more in-depth phenotypic data might ameliorate prediction models so that they can aid in informed decision-making for personalized medicine in cancer surgery. Previous prediction models have used albumin and hemoglobin levels, and it has been shown that a combination of physiological measurements and medical history makes better predictions [23, 33]. Adding preoperative laboratory data and past medical history to our prediction model might improve predictive performance. External validation will indicate the extent that the models can be generalized on populations outside of Denmark, and after clinical validation studies we believe that prediction modeling will be an integral part of the MDT in future.

## Conclusions

In this study, we demonstrated that it was possible to train and validate prediction models that could predict the occurrence of postoperative complications and anastomotic leakage. Discriminative power was deemed insufficient for current use in a clinical setting, but increasing data coverage to include a wider spectrum of the patient’s phenotype and genotype or constructing procedure-specific models, could improve models to become valuable for clinical practice.

Machine learning methods using observational health data utilized known factors associated with a higher risk for complications and anastomotic leakage as well as identified additional predictors, confirming the usefulness for a data-driven approach in prediction modeling.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10151-022-02624-x>.

**Acknowledgements** The Center for Surgical Science thanks European Health Data Evidence Network (EHDEN), EdenceHealth and Computerome (Danish Technical University) for their work contributing to this research project.

**Author contributions** VL, MG, AWR, KBB, JSRC, AT, EA, JSW, PR, ID and IG: made substantial contributions to the conception and design of the work as well as the acquisition, analysis, and interpretation of data for the work. VL, MG, AWR, KBB and JSRC: drafted the manuscript and AT, EA, JSW, PR, ID and IG: revised it critically for

important intellectual content. VL, MG, AWR, KBB, JSRC, AT, EA, JSW, PR, ID and IG: have given final approval of the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Funding** Innovative Medicines Initiative, 806968, Ismail Gögenur, Region Sjælland, Uddannelses- og Forskningsministeriet.

**Access to data** Data are owned by the Danish Colorectal Cancer Group (DCCG) and cannot be shared due to legislative restrictions. Parties interested in external validation can receive the necessary coding to test the prediction models on their own datasets if data has been converted to the OMOP CDM, without sharing sensitive data.

## Declarations

**Conflict of interest** Peter Rijnbeek has received a grant from the Innovative Medicines Initiative and consultancy fees from the Center for Surgical Science. The other authors declare that they have no conflict of interest.

**Ethics approval** This is an observational study for which no ethical approval or individual informed consent is required as determined by the National Committee on Health Research Ethics in Denmark.

## References

1. Dekker JWT, van den Broek CBM, Bastiaannet E, van de Geest LGM, Tollenaar RAEM, Liefers GJ (2011) Importance of the first postoperative year in the prognosis of elderly colorectal cancer patients. *Ann Surg Oncol* 18:1533–1539. <https://doi.org/10.1245/s10434-011-1671-x>
2. Duraes LC, Stocchi L, Steele SR et al (2018) The relationship between clavier-dindo morbidity classification and oncologic outcomes after colorectal cancer resection. *Ann Surg Oncol* 25:188–196. <https://doi.org/10.1245/s10434-017-6142-6>
3. Krarup P-M, Nordholm-Carstensen A, Jorgensen LN, Harling H (2014) Anastomotic leak increases distant recurrence and long-term mortality after curative resection for colonic cancer: a nationwide cohort study. *Ann Surg* 259:930–938. <https://doi.org/10.1097/SLA.0b013e3182a6f2fc>
4. Lee CHA, Kong JCH, Heriot AG, Warriar S, Zalberg J, Sitzler P (2019) Short-term outcome of emergency colorectal cancer surgery: results from Bi-National Colorectal Cancer Audit. *Int J Colorectal Dis* 34:63–69. <https://doi.org/10.1007/s00384-018-3169-5>
5. Frasson M, Granero-Castro P, Ramos Rodríguez JL et al (2016) Risk factors for anastomotic leak and postoperative morbidity and mortality after elective right colectomy for cancer: results from a prospective, multicentric study of 1102 patients. *Int J Colorectal Dis* 31:105–114. <https://doi.org/10.1007/s00384-015-2376-6>
6. Flynn DE, Mao D, Yerkovich ST et al (2020) The impact of comorbidities on post-operative complications following colorectal cancer surgery. *PLoS ONE* 15:e0243995. <https://doi.org/10.1371/journal.pone.0243995>
7. Merath K, Hyer JM, Mehta R et al (2020) Use of machine learning for prediction of patient risk of postoperative complications after liver, pancreatic, and colorectal surgery. *J Gastrointest Surg* 24:1843–1851. <https://doi.org/10.1007/s11605-019-04338-2>

8. Ingeholm P, Gögenur I, Iversen LH (2016) Danish colorectal cancer group database. *Clin Epidemiol* 8:465–468. <https://doi.org/10.2147/CLEP.S99481>
9. Klein MF, Gögenur I, Ingeholm P et al (2020) Validation of the Danish colorectal cancer group (DCCG.dk) database—on behalf of the Danish colorectal cancer group. *Colorectal Dis* 22:2057–2067. <https://doi.org/10.1111/codi.15352>
10. Clavien PA, Barkun J, de Oliveira ML et al (2009) The Clavien-Dindo classification of surgical complications: five-year experience. *Ann Surg* 250:187–196. <https://doi.org/10.1097/SLA.0b013e3181b13ca2>
11. Hripcsak G, Duke JD, Shah NH et al (2015) Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform* 216:574–578
12. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE (2012) Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc* 19:54–60. <https://doi.org/10.1136/amiainl-2011-000376>
13. Collins GS, Reitsma JB, Altman DG, Moons KGM (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 350:g7594. <https://doi.org/10.1136/bmj.g7594>
14. Rahbari NN, Weitz J, Hohenberger W et al (2010) Definition and grading of anastomotic leakage following anterior resection of the rectum: a proposal by the International Study Group of Rectal Cancer. *Surgery* 147:339–351. <https://doi.org/10.1016/j.surg.2009.10.012>
15. Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR (2018) Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *J Am Med Inform Assoc* 25:969–975. <https://doi.org/10.1093/jamia/ocy032>
16. Khalid S, Yang C, Blacketer C et al (2021) A standardized analytics pipeline for reliable and rapid development and validation of prediction models using observational health data. *Comput Methods Programs Biomed* 211:106394. <https://doi.org/10.1016/j.cmpb.2021.106394>
17. Frasson M, Flor-Lorente B, Rodríguez JLR et al (2015) Risk factors for anastomotic leak after colon resection for cancer: multivariate analysis and nomogram from a multicentric, prospective, national study with 3193 patients. *Ann Surg* 262:321–330. <https://doi.org/10.1097/SLA.0000000000000973>
18. Bilimoria KY, Liu Y, Paruch JL et al (2013) 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* 217:833–842.e3. <https://doi.org/10.1016/j.jamcollsurg.2013.07.385>
19. Keller DS, Ho JW, Mercadel AJ, Ogola GO, Steele SR (2018) Are we taking a risk with risk assessment tools? Evaluating the relationship between NSQIP and the ACS risk calculator in colorectal surgery. *Am J Surg* 216:645–651. <https://doi.org/10.1016/j.amjsurg.2018.07.015>
20. Ziegler MA, Bauman JC, Welsh RJ, Wasvary HJ (2020) Can the American college of surgeons national surgical quality improvement program risk calculator predict outcomes for urgent colectomies? *Am Surg*. <https://doi.org/10.1177/0003134820973392>
21. Ma M, Liu Y, Gotoh M et al (2021) Validation study of the ACS NSQIP surgical risk calculator for two procedures in Japan. *Am J Surg*. <https://doi.org/10.1016/j.amjsurg.2021.06.008>
22. Sherman SK, Hrabe JE, Huang E, Cromwell JW, Byrn JC (2018) Prospective validation of the Iowa rectal surgery risk calculator. *J Gastrointest* 22:1258–1267. <https://doi.org/10.1007/s11605-018-3770-5>
23. Watanabe T, Miyata H, Konno H et al (2017) Prediction model for complications after low anterior resection based on data from 33,411 Japanese patients included in the national clinical database. *Surgery* 161:1597–1608. <https://doi.org/10.1016/j.surg.2016.12.011>
24. Arezzo A, Migliore M, Chiaro P et al (2019) The REAL (Rectal anastomotic leak) score for prediction of anastomotic leak after rectal cancer surgery. *Tech Coloproctol* 23:649–663. <https://doi.org/10.1007/s10151-019-02028-4>
25. Dekker JWT, Liefers GJ, de van Mol Otterloo JCA, Putter H, Tollenaar RAEM (2011) Predicting the risk of anastomotic leakage in left-sided colorectal surgery using a colon leakage score. *J Surg Res* 166:27–34. <https://doi.org/10.1016/j.jss.2010.11.004>
26. Sammour T, Cohen L, Karunatilake AI et al (2017) Validation of an online risk calculator for the prediction of anastomotic leak after colon cancer surgery and preliminary exploration of artificial intelligence-based analytics. *Tech Coloproctol* 21:869–877. <https://doi.org/10.1007/s10151-017-1701-1>
27. Klose J, Tarantino I, von Fournier A et al (2018) A nomogram to predict anastomotic leakage in open rectal surgery—hope or hype? *J Gastrointest* 22:1619–1630. <https://doi.org/10.1007/s11605-018-3782-1>
28. Yang SU, Park EJ, Baik SH, Lee KY, Kang J (2019) Modified colon leakage score to predict anastomotic leakage in patients who underwent left-sided colorectal surgery. *J Clin Med*. <https://doi.org/10.3390/jcm8091450>
29. Yu X-Q, Zhao B, Zhou W-P et al (2016) Utility of colon leakage score in left-sided colorectal surgery. *J Surg Res* 202:398–402. <https://doi.org/10.1016/j.jss.2015.12.046>
30. Marres CCM, van de Ven AWH, Verbeek PCM, van Dieren S, Bemelman WA, Buskens CJ (2016) The effect of a postoperative quality improvement program on outcomes in colorectal surgery in a community hospital. *Int J Colorectal Dis* 31:1603–1609. <https://doi.org/10.1007/s00384-016-2619-1>
31. Carli F, Charlebois P, Baldini G, Cachero O, Stein B (2009) An integrated multidisciplinary approach to implementation of a fast-track program for laparoscopic colorectal surgery. *Can J Anaesth* 56:837–842. <https://doi.org/10.1007/s12630-009-9159-x>
32. Trépanier M, Minnella EM, Paradis T et al (2019) Improved disease-free survival after prehabilitation for colorectal cancer surgery. *Ann Surg* 270:493–501. <https://doi.org/10.1097/SLA.0000000000003465>
33. Nielsen AB, Thorsen-Meyer H-C, Belling K et al (2019) Survival prediction in intensive-care units based on aggregation of long-term disease history and acute physiology: a retrospective study of the Danish national patient registry and electronic patient records. *Lancet Digit Health* 1:e78–e89. [https://doi.org/10.1016/S2589-7500\(19\)30024-X](https://doi.org/10.1016/S2589-7500(19)30024-X)

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