Systematic review of preoperative and intraoperative colorectal Anastomotic Leak Prediction Scores (ALPS)

Venn, Mary L; Hooper, Richard L; Pampiglione, Tom; Morton, Dion G; Nepogodiev, Dmitri; Knowles, Charles H

DOI:
10.1136/bmjopen-2023-073085

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Venn, ML, Hooper, RL, Pampiglione, T, Morton, DG, Nepogodiev, D & Knowles, CH 2023, 'Systematic review of preoperative and intraoperative colorectal Anastomotic Leak Prediction Scores (ALPS)', *BMJ open*, vol. 13, no. 7, e073085. https://doi.org/10.1136/bmjopen-2023-073085

BMJ Open

Systematic review of preoperative and intraoperative colorectal Anastomotic Leak Prediction Scores (ALPS)

Mary L Venn [1], Richard L Hooper [2], Tom Pampiglione [1], Dion G Morton [3], Dmitri Nepogodiev [3], Charles H Knowles [1]

To cite: Venn ML, Hooper RL, Pampiglione T, *et al*. Systematic review of preoperative and intraoperative colorectal Anastomotic Leak Prediction Scores (ALPS). *BMJ Open* 2023;**13**:e073085. doi:10.1136/bmjopen-2023-073085

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2023-073085).

Received 28 February 2023
Accepted 26 June 2023

[1]Blizard Institute, Queen Mary University of London, London, UK
[2]Institute of Population Health Sciences, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
[3]NIHR Global Health Research Unit on Global Surgery, Institute of Translational Medicine, University of Birmingham Edgbaston Campus, Birmingham, UK

Correspondence to
Mary L Venn;
maryvenn@doctors.org.uk

ABSTRACT

Objective To systematically review preoperative and intraoperative Anastomotic Leak Prediction Scores (ALPS) and validation studies to evaluate performance and utility in surgical decision-making. Anastomotic leak (AL) is the most feared complication of colorectal surgery. Individualised leak risk could guide anastomosis and/or diverting stoma.

Methods Systematic search of Ovid MEDLINE and Embase databases, 30 October 2020, identified existing ALPS and validation studies. All records including >1 risk factor, used to develop new, or to validate existing models for preoperative or intraoperative use to predict colorectal AL, were selected. Data extraction followed CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies guidelines. Models were assessed for applicability for surgical decision-making and risk of bias using Prediction model Risk Of Bias ASsessment Tool.

Results 34 studies were identified containing 31 individual ALPS (12 colonic/colorectal, 19 rectal) and 6 papers with validation studies only. Development dataset patient populations were heterogeneous in terms of numbers, indication for surgery, urgency and stoma inclusion. Heterogeneity precluded meta-analysis. Definitions and timeframe for AL were available in only 22 and 11 ALPS, respectively. 26/31 studies used some form of multivariable logistic regression in their modelling. Models included 3–33 individual predictors. 27/31 studies reported model discrimination performance but just 18/31 reported calibration. 15/31 ALPS were reported with external validation, 9/31 with internal validation alone and 4 published without any validation. 27/31 ALPS and every validation study were scored high risk of bias in model analysis.

Conclusions Poor reporting practices and methodological shortcomings limit wider adoption of published ALPS. Several models appear to perform well in discriminating patients at highest AL risk but all raise concerns over risk of bias, and nearly all over wider applicability. Large-scale, precisely reported external validation studies are required.

PROSPERO registration number CRD42020164804.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Rigorous systematic methodology: inclusion criteria unlikely to have missed further Anastomotic Leak Prediction Scores.
⇒ Rigorous methodological evaluation of model construction, applicability and risk of bias.
⇒ Exclusion of prediction scores using postoperatively assessed factors.
⇒ No individual participant data meta-analysis was undertaken due to wide heterogeneity.
⇒ Unable to draw clinical conclusions due to poor overall methodological quality.

INTRODUCTION

Anastomotic leak (AL) is the most serious complication following colorectal surgery causing significant morbidity and mortality.[1] Incidences in the 2015 and 2017 European Society of Coloproctology international audits of right and left sided resections were 8.1%[2] and 8.6%,[3] respectively. Several factors were identified in univariate analyses that increased risk of AL after right-sided resection including patient factors known preoperatively such as gender, indication, operative urgency and smoking status; and operative factors such as approach (open vs laparoscopic).[2] In left-sided resections, ALs were more common with male sex, neo-adjuvant treatment, more distal anastomosis, hand-sewn anastomosis, defunctioning ileostomy or planned postoperative critical care admission.[4] Such data indicate that careful patient selection has a role in reducing the risk of AL.

Despite a significant body of evidence to describe risk factors for AL, and a multitude of published risk scores, there is no current consensus among the surgical community about the best prediction model to stratify patients preoperatively or intraoperatively for AL risk. It has been clearly demonstrated that surgeon estimates are poor[5 6] making objective clinical measures a priority.

While AL is perilous, overuse of diverting ileostomy also creates morbidity, stoma complications, reduces quality-of-life, causes additional costs and need for further surgery. Better risk stratification on an individual

patient basis could translate to reduced morbidity and mortality as well as associated cost savings. Individually calculated risk assessment would aid informed shared decision-making between patients and surgeons.

Study aim

To systematically review existing preoperative and intraoperative colorectal AL prediction models, evaluate their performance including any validation studies and assess their ability to aide surgical decision making.

METHODS

This systematic review is reported according to The CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (online supplemental CHARMS checklist),[7] to ensure appraisal of predictive models was performed rigorously and reproducibly. The study goes beyond the protocol to also report calibration (as well as planned discrimination, validation and quality assessments). Preferred Reporting Items for Systematic Reviews and Meta-Analyses[8] checklist is included in the supplementary material.

Data sources

A literature search was performed in Ovid MEDLINE and Embase electronic databases to identify studies between January 1990 to the search date containing keywords 'colorectal' or derivatives thereof; 'anastomosis/anastomotic'; 'leak' or 'breakdown' or 'failure' or 'dehiscence' and 'tool' or 'model' or 'nomogram' or 'risk score', see supplementary material for full search details. Reference lists of included studies were searched for further studies. No language restrictions were set. Searches were completed in Ovid MEDLINE 30 October 2020 and Embase 28 October 2020.

Study selection

Inclusion criteria: all studies proposing and/or validating a predictive model that could be used preoperatively or intraoperatively to predict AL following colonic or rectal resection. To qualify as a score or model, more than one predictive factor (variable) must be used in the score. Studies that evaluated risk factors but neither generated nor validated a risk score/model/nomogram were excluded. Studies that reported a model that required any variable to be obtained postoperatively were excluded. Abstracts that contained sufficient data were included where full text was unavailable.

All titles and abstracts obtained in the searches were screened for inclusion. Non-human studies, non-colorectal surgery studies, those examining a different outcome or not proposing a predictive score that is, single factor studies, were excluded. Abstracts were screened manually (in Microsoft Excel) by a single author, except where doubt arose and CK was consulted. Articles included based on abstract review were imported to Mendeley Desktop V.1.19.8 for full text review. Screening was not blinded, conflicts and uncertainties were resolved through discussion. Full texts were obtained for relevant articles (where available).

Data extraction

The CHARMS checklist[7] was used for complete data extraction including data source and dates for each study, participants, outcome definitions, variables (predictors), model development method and model performance. Both apparent performance (performance of the score in the model development set) and any internal or external validation methods and results were extracted. A pragmatic approach was taken to extract and analyse any measures of model performance such as discrimination (including the area under the (receiver operating characteristic) curve (AUC)), calibration, classification or overall performance measures. Two independent reviewers (MLV, TP) extracted variables from papers in duplicate. Any disagreement was resolved by discussion with the senior author.

Quality assessment

Each study was assessed for applicability and risk of bias by two independent reviewers (MLV and TP). Any discrepancies were resolved by discussion. Applicability considered the extent to which the model could be used in preoperative or intraoperative prediction of AL according to CHARMS criteria; it is reported as a 'concern level'. Applicability criteria (see table 1) were adapted from a framework used in a systematic review of prediction models of the outcomes of colorectal cancer in patients ≥65 years.[9]

PROBAST criteria (Prediction model Risk Of Bias ASsessment Tool)[10 11] were used to assess risk of bias in four domains (participants, predictors, outcome and analysis) that can introduce systematic bias to model performance calculations. Risk of bias questions were tailored for use with model development *or* validation studies and are detailed in online supplemental table S1.

Data synthesis and analysis

Extracted data are reported in tables including study design, modelling methods, patient characteristics and outcome definitions; predictors and presentation of final model (ie, score chart, nomogram or calculator (paper or online)); and model performance with associated validation studies. Quality assessments (applicability and risk of bias) are described for all models. Online supplemental material: predictor selection provides further detail about statistical methods for predictor selection and modelling methods used for each model developed.

Patient and public involvement

No patient involved.

RESULTS

Search findings

The literature search identified 834 records from OVID and Embase with 8 additional records identified through

other means, for example, searches of reference lists. Seven hundred and two records remained after de-duplication and 642 were excluded on screening. Sixty publications were further reviewed for inclusion (figure 1), and yielded 34 records. Of these, 28 records included 31 separate Anastomotic Leak Prediction Scores (ALPS) (with or without validation studies) and six records contained external validation studies of prediction models (without model development or adaptation) (see online supplemental tables S2a and S2ab).

For three study abstracts, full text English-language articles were unavailable but adequate information was presented to merit inclusion in the review; Jiang *et al*[12] published only Chinese full text; McKenna *et al*[13] published a plenary presentation abstract; and Yao *et al*[14] was out of print.

Explanatory note

One ALPS that was externally validated by Sammour *et al*[15] is the ACS NSQIP (American College of Surgeons National Surgical Quality Improvement Programme) online calculator that may be found at riskcalculator.facs.org. Unfortunately, although the online calculator offers risk prediction for AL, the Surgical Risk Calculator development paper[16] did not document AL as a defined outcome nor did it list the factors, method or model used to offer AL risk prediction, either online or in published material. As such it has not been included in the main

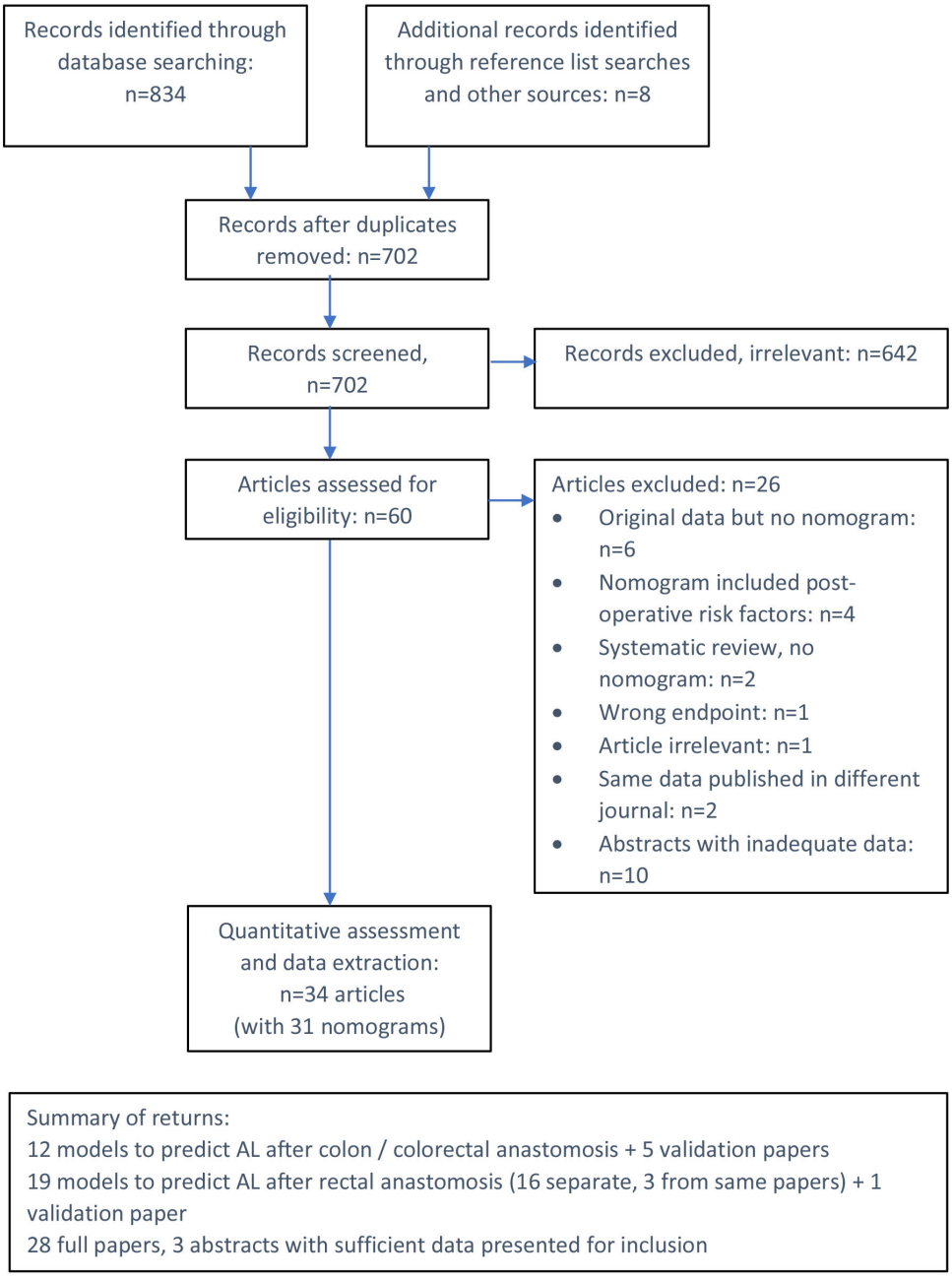

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram.[8] AL, anastomotic leak.

results. Of note, in Sammour's study the ACS NSQIP online score performed poorly, generating AUC of only 0.58.[15] ACS NSQIP registry data have however been used to develop three other models that are included in this review.[13 17 18]

Study designs

Included records are summarised in online supplemental tables S2a and S2b. These outline the study origin, key details of development datasets and patient/surgical characteristics. They also report how the outcome (AL) was defined, and identify the model validation methods presented. For ease of comparison, studies are divided into ALPS for colonic or mixed colorectal anastomoses (online supplemental table S2a) versus those for rectal anastomosis (online supplemental table S2b). All results are divided into these two categories and presented by date of publication. The earliest preoperative/intraoperative ALPS was published in 2011. There was a breadth of geographical study origin from East Asia, Europe and the USA and validation studies from Australia and South America.

Methods for model development were heterogeneous with multivariable logistic regression (MVLR) featuring most frequently (8/12 colonic/colorectal ALPS and 18/19 rectal ALPS) with or without methods to enhance calibration such as ridge regression or LASSO (least absolute shrinkage and selection operator) (4/31 ALPS)[19–21]; these methods are employed to reduce model overfit.[22 23] Other authors employed: systematic review methodology with meta-analysis of risk factors (using individual patient data[24] or pooling ORs from individual studies[25]); systematic review with Delphi consensus for factor value (1/31)[26]; or machine learning (2/31).[27 28] Model development datasets were also heterogeneous with single site data (15/31), multicentre studies (6/31), registry data (5/31) or data from systematic review creating the ALPS development dataset (4/31), in one, the source was not recorded. The number of patients included in model development ranged from 79[29] to 37 950.[18]

Patient populations

There was considerable heterogeneity between patient populations, see online supplemental tables S2a and S2b. All development datasets included patients undergoing cancer resection but studies that used registry data such as the ACS NSQIP data[13 18] or International TaTME Registry[30] as well as those using systematic review modelling method also included patients undergoing benign bowel resections. While most studies included any adult patients, two developed models exclusively for patients ≥65 years.[17 31] Among patients with colonic or colorectal resections, 4/12 models included non-elective surgery compared with 2/19 models to detect AL after rectal resection. Assessment of operative approach, that is, whether open, laparoscopic or any mode surgery was included; and inclusion (or exclusion) of patients with a diverting stoma revealed notable differences between the developed models. Patients with diverting stoma were specifically included in development datasets for 3/12 colonic/colorectal and 11/19 rectal resection AL prediction models.

Outcomes

Of all 31 models developed, AL was defined in 22 studies, undefined in nine. Twenty models did not record the outcome timeframe, and of the 11 that did, it ranged from 'during index admission'[32] to 3 months[33 34] or 'no time limit'.[35] There was further variability between the studies in AL definition, with some studies including only AL that required some mode of treatment, and others including all cases of apparent AL on imaging regardless of clinical course.

Validation

To validate their ALPS, authors used a variety of internal or external validation methods. The validation methods may be ranked for their risk of bias or overfitting according to online supplemental figure S1, with the apparent performance (performance estimated from the same dataset used to develop the model) at highest risk of bias. Of colonic/colorectal ALPS, 9/12 attempted some form of external validation even if only in a subset of patients with diverting stoma, the remaining three studies reported internal split sample or 10-fold cross validation. In contrast, only 6/19 rectal ALPS described any attempt at external validation, 9/19 reported internal validation using split sample or bootstrap techniques and four studies made no attempt to validate their novel model. TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) provides explanations that describe the differences between these methods and advantages/drawbacks of different methods.[36]

Predictors

The number of predictors in the models varied between 3 and 33.[19] All predictive factors used in the different ALPS are displayed in online supplemental tables S3a and S3b. The single most frequent predictor was patient sex, used in 7/12 and 14/19 colorectal/rectal ALPS, respectively.

Of colonic/colorectal ALPS, 7/12 used American Society of Anesthesiologists grade (ASA); 6/12 used toxins (smoking, alcohol or steroids), 4/12 used body mass index (BMI), 4/12 used neoadjuvant treatments, 3/12 used diabetes and 3/12 used age. 7/12 colonic/colorectal ALPS required blood test results and 8/12 models required intraoperative details to complete the score. The two models from Soguero-Ruiz *et al* required electronic health record (EHR) data and machine learning techniques.[27 28]

Of the 19 rectal ALPS, 8/19 used diabetes, 5/19 BMI, 4/19 ASA and just 2/19 age. 6/19 rectal ALPS featured toxins; all six used smoking, with alcohol in 3/19 and steroids as additional risk factor in 1/19. Neoadjuvant treatments were a factor in 9/19 models. 9/19 rectal ALPS

required blood test results with albumin level featuring seven times. 16/19 scores could not be calculated until intraoperative information was available and 15/19 scores relied on tumour information, for example, distance to anus or tumour diameter. Two rectal ALPS required specialist input: Xiao *et al*'s ALPS used measurement of microvascular density in distal margin,[35] made possible only with pathology expertise on-hand intraoperatively; Yu *et al*'s study[33] used measurement of pelvic dimensions. While this can be easily calculated with modern CT equipment, the surgeon must have forethought to review in conjunction with the radiologist.

Predictor selection

Generally, studies either selected predictors based on univariate analysis then performed MVLR, then ascribing the corresponding weighting of each independent risk factor to produce a prediction model, or used MVLR to identify significant risk factors using another approach such as stepwise selection. Methods for predictor selection are described in supplementary material: Predictor selection. Three authors (producing four ALPS) recognised the risk of overfitting their models and attempted to control this with ridge or LASSO regression techniques.[19–21]

One model[26] used Delphi consensus to assign the weight of each risk factor identified at systematic review. Soguero-Ruiz *et al* used the free text from EHRs to create a prediction model[27] and augmented this with the addition of blood results and vital signs[28] in a second published model.

Applicability

Table 1 shows the 'concern levels' for model applicability according to the CHARMS checklist,[7] where models are assessed for their ability to preoperative or intraoperatively predict AL following colorectal resection and primary anastomosis. 5/12 colonic/colorectal and 13/19 rectal ALPS attracted low concern for applicability in *participant selection*. Concerns were raised when inclusion criteria were undefined,[27] a small cohort was used, or key information (eg, stoma inclusion) was missing, but only reached high level of concern in one study that published few details about the development dataset.[27]

Studies with *predictors* that are readily available preoperatively attracted low applicability concerns however *all* colonic/colorectal ALPS and 16/19 rectal ALPS had at least moderate concern due to the need for either intraoperative information (eg, intraoperative blood loss or transfusion, height of anastomosis, operation time); requirement for specialist assessment (eg, pelvic dimensions on imaging[33] or machine learning techniques[27 28]); or >20 factors included.[19 25] 3/19 rectal ALPS achieved low applicability concern for predictors[12 30 37] with an otherwise almost even mix of moderate and high concern level.

Applicability concern was low for *outcomes* in 3/12 colonic/colorectal ALPS and 7/19 rectal ALPS. This could be achieved if the study defined both the AL and the prediction horizon (timeline to detect AL).[38] 5/12 and 3/19 colonic/colorectal and rectal ALPS neither defined AL nor the timeframe in which the complication was sought, inferring high concern level.

Model performance

Online supplemental tables S4a and S4b show that the most common reported performance measure was the AUC, reported for 27/31 ALPS. This is a measure of discrimination and with a binary endpoint, is equivalent to the concordance or C-index reported in some studies and marked as '©' in the table. The measure represents the probability that for any random pair of individuals, one with, one without the outcome, the model will assign a higher probability to the one *with* the outcome. An AUC or C-index was reported for 27/31 ALPS models, either measuring the apparent performance only[17 24 33 39] (performance in development dataset), or reporting an internal or external validation test.

Among colonic/colorectal ALPS, Pasic and Salkic[29] reported the highest AUC of 1.0 though they used the smallest number of patients (n=40) in a split sample internal validation cohort. The lowest reported AUC of 0.62 was reported by Frasson *et al*,[40] using a 10-fold internal cross-validation dataset. For rectal ALPS, Crispin *et al*,[19] documented the weakest discrimination at 0.595 while Cheng *et al*,[21] achieved an AUC of 0.952 in a small (n=94) split sample internal validation cohort.

Seven studies performed independent external validation, that is, validation of an ALPS other than their own model. Interestingly, only two ALPS; CLS score (Colon Leakage Score)[26] and ANACO (ANAstomotic leak after COlon resection for cancer)[40] were subject to independent external validation from a separate research group. Five of seven papers reporting evaluation of another ALPS were *only* validation papers (see external validation papers in online supplemental table S2a) with others[20 25] re-evaluating CLS ALPS[26] in addition to developing their own ALPS. The results from external validation papers are reported in bold (online supplemental table S4) with the original published model shown underneath for comparison. The external validation datasets included fewer patients (range 83–972 patients) and consequently had fewer events than the model development studies. These smaller cohorts result in weaker discrimination results.

In addition to discrimination, 5/12 colonic/colorectal ALPS (only three in external validation datasets) and 13/19 rectal ALPS presented results for calibration. For rectal ALPS, three were calibrated in the original unadjusted dataset, six in internal validation and only four models in external validation datasets. Models tend to be well-calibrated in their own development dataset, which diminishes the value of calibration results in this context. Calibration was usually reported in plots or charts by comparing predicted risk by deciles, against observed AL rate. Three groups[12 25 31] used a Hosmer-Lemeshow

Table 1 Applicability concern levels of Anastomotic Leak Prediction Scores models according to CHARMS checklist

Colon/colorectal resection models	Applicability concern level			Rectal resection models	Applicability concern level		
Author, score	Participant selection	Predictors	Outcome	Author, score	Participant selection	Predictors	Outcome
Dekker *et al*[26], CLS	M	M	M	Park *et al*[43]	L	M	L
Pasic and Salkic[29]	M	M	M	Yao *et al*[14]	M	M	M
Frasson *et al*[40], ANACO	L	M	L	Liu *et al*[51]	L	H	M
Rojas-Machado *et al*[25], PROCOLE IP	M	H	M	Hu and Cheng[32]	L	H	L
Soguero-Ruiz *et al*[27], BoW	H	H	H	Rojas-Machado *et al*[25], PROCOLE LAR*	M	H	M
Soguero-Ruiz *et al*[28], Hetero-data	M	H	H	Crispin *et al*[19], DGAV rectum*	M	H	H
Crispin *et al*[19], DGAV colon	M	H	H	Kim *et al*[52]	L	M	M
Rencuzogullari *et al*[17], ACS NSQIP>65	L	M	L	Watanabe *et al*[37]	L	L	M
Shen *et al*[31]	M	M	H	Hoshino *et al*[46]	L	M	M
Yang *et al*[20], m-CLS	L	M	M	Liu *et al*[34]	M	M	L
McKenna *et al*[18], LEFT ACS NSQIP	L	M	L	Cheng *et al*[21]	L	M	M
McKenna *et al*[13], RIGHT ACS NSQIP	L	M	H	Arezzo *et al*[24]	M	H	H
Table does not include validation studies.				Jiang *et al*[12]	M	L	H
				Penna *et al*[30]	L	L	L
				Xiao *et al*[35]	L	H	L
				Zheng *et al*[47]	L	M	M
				Shiwakoti *et al*[39]	L	M	M
Key to applicability scores described below, developed according to CHARMS criteria[7]:				Yu *et al*[33], Gender	L	M	L
				*Yu *et al*[33], Pelvic dimensions	L	H	L

Aspect of model development		Grade	Criteria (AL)
		Applicability concern …	
Participant selection	L	Low if …	Consecutive participants, well defined inclusion/exclusion criteria, ideally multiple settings/centres
	M	Moderate if …	Somewhat representative of colorectal population, mostly defined, some information absent (eg, whether stoma included) or small cohort
	H	High if …	Not representative of colorectal surgery population OR inclusion criteria missing
		Applicability concern …	
Predictors	L	Low if …	Predictors readily available for all patients AND available preoperatively
	M	Moderate if …	Predictors readily available for all patients but may include intraoperative factors OR available preoperatively but one or more factor may not be universally available

Continued

Table 1 Continued

Aspect of model development		Grade	Criteria (AL)
	H	High if …	Neither criteria met, or high inter-operator variability of factor for example, height of anastomosis, or specialist assessment required (radiology, histopathology), or >20 factors
		Applicability concern …	
Outcome	L	Low if …	AL AND duration of follow-up for AL defined
	M	Moderate if …	AL OR duration of follow-up for AL not defined
	H	High if …	Neither AL criteria met

*Same paper, additional prediction model.
ACS-NSQIP, American College of Surgeons National Surgical Quality Improvement Programme; AL, anastomotic leak; ANACO, Spanish study on ANAstomotic leak after COlon resection for cancer; BoW, bag of words (term used for free text extracted from electronic health records); CHARMS, CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies; CLS, Colon Leakage Score; DGAV, German Society for General and Visceral Surgery; m-CLS, modified Colon Leakage Score; PROCOLE IP, prognostic colorectal leakage score intra-peritoneal; PROCOLE LAR, prognostic colorectal leakage score low anterior resection.

test of agreement that divides patients into 10 groups by predicted probabilities and computes a χ^2 statistic from the observed versus expected frequencies. Dekker *et al*[26] produced a scatter plot of the CLS score in consecutive patients, colouring patients with AL a different colour but this is not a true calibration plot.

When a probability threshold is selected, it is also possible to report classification measures such as sensitivity, specificity, positive and negative predictive values and likelihood ratios all of which have been displayed when reported (5/12 colonic/colorectal and 2/19 rectal ALPS). Finally, overall performance measures may be employed, for example, Crispin *et al* used the Brier score, but did not report discrimination.[19] The Brier score[36 41] evaluates accuracy of prediction. A lower score (near 0) indicates a better performance. It is not considered a good test for prediction models with imbalanced classes so may be considered an inferior choice of test in predicting AL in cohort groups. Moreover, the Brier score was only used in one study in this review so could not be used as a model comparator.

Three Chinese research groups[33–35] attempted to ascertain the clinical value of their model using 'decision curve analysis' to obtain a value of net benefit, this is a figure calculated based on the true positive minus the false positive rate of using the model at a particular threshold.[42] Pasic and Salkic[29] tested their model in a small cohort of patients to evaluate its effectiveness in avoiding AL by altering the operative plan at particular risk thresholds, see online supplemental file 1 case study 1.[29] Park *et al*[43] re-analysed patients excluded from the development dataset due to diverting stoma placement and calculated the number of stomas that would have been avoided were the model used, see online supplemental file 1 case study 2.

The heterogeneity in patient groups, outcomes and performance reporting made meta-analysis of the performance of predictors in our review impossible. Examples of ROC curves, calibration plots and nomograms can be found in TRIPOD pages W51–W53.

Risk of bias

Every paper was dual assessed, see table 2, for risk of bias in the following domains; participant, predictor, outcome and analysis. Results for colonic/colorectal and rectal ALPS were similar. Risk of bias in *participant* selection was low in 8/12 colonic/colorectal and 17/19 rectal ALPS but increased where patients were selected in an unsystematic manner[27 28] or patients with missing variables were excluded.[31] The exclusion of patients with missing (variable) data repeatedly increased the risk of bias in the external validation studies.[6 15 44 45]

Predictors were generally well-defined and were available for use preoperatively or intraoperatively leading to low risk of bias in 10/12 colonic/colorectal and 9/19 rectal ALPS. By contrast, risk of bias was raised where the model was developed through systematic review such that some variables were undefined,[24 26] or if predictors were difficult to obtain intraoperatively such as microvascular density in the distal resection margin[35] or left colic artery preservation status that is subject to error.[14] Inclusion of variables such as tumour diameter[30 33 39 46 47] that could be subject to error unless the resected specimens accurately measured, introduced unclear risk of bias.

Where AL was undefined, the risk of bias for *outcomes* was recorded as high; where the time interval for AL to occur was not documented, this was recorded as unclear risk of bias. Increased risk of bias was also introduced where the outcome was defined in different ways for some participants, for example, in datasets derived from

Table 2 Risk of bias level according to PROBAST (Prediction model Risk Of Bias ASsessment Tool)[10 11]

Colon/colorectal resection models	Potential bias			
Author, model	**Data source and participants**	**Predictors**	**Outcome**	**Analysis**
Dekker *et al*[26], CLS	Unclear	High	High	High
Pasic and Salkic[29]	Low	Low	Unclear	High
Frasson *et al*[40], ANACO	Low	Low	Low	High
Rojas-Machado *et al*[25], PROCOLE IP	Low	Unclear	Unclear	High
Soguero-Ruiz *et al*[27], BoW	High	Low	High	High
Soguero-Ruiz *et al*[28], Hetero-data	High	Low	High	High
Crispin *et al*[19], DGAV colon	Low	Low	High	Low
Rencuzogullari *et al*[17], ACS NSQIP>65	Low	Low	Low	High
Shen *et al*[31]	High	Low	High	High
Yang *et al*[20], m-CLS	Low	Low	Unclear	High
McKenna *et al*[18], LEFT ACS NSQIP	Low	Low	Low	High
McKenna *et al*[13], RIGHT ACS NSQIP #	Low	Low	High	High
Validation studies				
Yu *et al*[44]	High	Low	Unclear	High
Sammour *et al*[6]	High	Low	Low	High
Sammour *et al*[15]	High	Low	Low	High
Muñoz *et al*[45]	High	Low	Unclear	High
Klose *et al*[53]	High	Unclear	Unclear	High
Rectal resection models	**Potential bias**			
Author, model	**Data source and participants**	**Predictors**	**Outcome**	**Analysis**
Park *et al*[43]	Low	Low	Low	High
Yao *et al*[14]	Unclear	High	High	Unclear
Liu Y *et al*[51]	Low	Low	Unclear	High
Hu *et al*[32]	Low	Low	Low	High
Rojas-Machado *et al*[25], PROCOLE LAR*	Low	Unclear	Unclear	High
Crispin *et al*[19], DGAV rectum*	Low	Low	High	Low
Kim *et al*[52]	Low	Low	Unclear	High
Watanabe *et al*[37]	Low	Low	Unclear	High
Hoshino *et al*[46]	Low	Unclear	Unclear	High
Liu *et al*[34]	Low	Low	Low	High
Cheng *et al*[21]	Low	Low	Unclear	High
Arezzo *et al*[24]	Unclear	High	High	High
Jiang *et al*[12]†	Low	Low	High	High
Penna *et al*[30]	Low	Unclear	Low	High
Xiao *et al*[35]	Low	High	Low	High
Zheng *et al*[47]	Low	Unclear	Unclear	Low
Shiwakoti *et al*[39]	Low	Unclear	Low	High
Yu *et al*[33]	Low	Unclear	Low	High
Yu* *et al*[33]	Low	Unclear	Low	High
Validation studies				
Miyakita *et al*[37]	High	Unclear	Low	High

PROBAST criteria for applicability scores in online supplemental table 1.
*Additional model in same paper.
†Abstract only.
ACS-NSQIP, American College of Surgeons National Surgical Quality Improvement Programme; ANACO, Spanish study on ANAstomotic leak after COlon resection for cancer; BoW, bag of words (term used for free text extracted from electronic health records); CLS, Colon Leakage Score; DGAV, German Society for General and Visceral Surgery; m-CLS, modified Colon Leakage Score; PROCOLE IP, prognostic colorectal leakage score intra-peritoneal; PROCOLE LAR, prognostic colorectal leakage score low anterior resection.

a systematic review. 6/12 colonic/colorectal and 4/19 rectal ALPS were classified as high risk of outcomes bias with another 4/12 and 6/19 at unclear risk of bias.

Risk of bias in *analysis* was assessed for the strongest form of validation reported in each paper that is, external validation if available, then internal validation or apparent performance if no validation tests were reported (see online supplemental figure S1 for validation hierarchy). Risk of bias in analysis was almost uniformly high (11/12 colonic/colorectal and 16/19 rectal ALPS) with a broad range of weaknesses including low event or event: variable rate; mismanagement of continuous predictors converted into ≥2 categories in prediction models; inadequate handling of missing data; and incomplete reporting of key model performance measures such as discrimination and calibration. Only two studies achieved low risk of analysis bias.[19 47] Among the external validation papers, all six were at high risk of bias with small cohorts and inadequate event: variable rates.

DISCUSSION

This review identified 31 ALPS in 28 model development studies; 12 for colonic or colorectal AL and 19 for rectal AL. There were methodological concerns in most studies, with 16 models lacking external validation, only 18 studies reporting calibration and, strikingly, a high risk of analysis bias identified in 27/31 models. The six studies offering independent external validation only centred around two models (CLS[26] and ANACO[40]). They tended to have small sample sizes, did not properly report model calibration and, despite some improved discrimination results, could not be used to support model use in clinical practice due to consistently high risk of participant *and* analysis bias. As per TRIPOD guidance, 'In validation studies, assessment of both discrimination and calibration is fundamental'.[36] All but one study by Penna *et al*,[30] raised concerns for applicability, putting into question how the ALPS could actually be applied to colorectal patients in practice. Models that are likely best calibrated are those that have used ridge or LASSO regression,[22 23] but without external validation, these models must still be used with caution.

Study limitations

Study selection was rigorous and it is unlikely we have missed further ALPS (beyond ACSNSQIP's online calculator, see the Results section and the Search findings subsection) however our review has some limitations. First, it included only preoperative and intraoperative ALPS on the basis that the best opportunity to avoid an anastomosis or choose a diverting stoma is in theatre. Other prediction scores, which may perform well (and have a place in ruling out leak to enable early postoperative discharge), were excluded for including postoperatively assessed factors. Second, we have not attempted an individual participant data meta-analysis using multiple datasets from our review, though such an approach might offer an opportunity for cross-validating existing models or developing a new model. Key challenges included the lack of uniformity in defining AL and the mixed policies for inclusion or exclusion of patients with diverting stomas.

Third, search strategies could be further optimised: our latest search is 2 years ago. Though this review searched two databases, recommendations from the Study Center of the German Society of Surgery, propose that surgical systematic reviews should in addition search Web of Science, and Cochrane Central Register of Controlled Trials (CENTRAL).[48] We also reflect that exact search terms should be recorded including controlled vocabulary; the use of the Boolean operator NOT for 'vascular anastomosis' is not considered best practice, and duplicate, blinded citation screening for *all* citations would strengthen the methodology.

Improving methodology

Reporting practice has changed since 2015 TRIPOD guidance.[36] Studies published prior to 2015 frequently lacked key performance measures that enable comparisons to be drawn. Where external validation studies were performed of older (pre-2015) ALPS, there was a tendency for validation study authors to report similar statistical measures and reproduce any miscellaneous figures published in the original model development study, adapted for the external cohort.[45] While model performance reporting before TRIPOD was variable, uniformity in reporting is still lacking, however, some papers have demonstrated better guidelines adherence.[33 47]

Risk of bias is a critical aspect of study evaluation but is frequently omitted from review papers or ignored by the reader. Where performance measures lead the reader to believe an ALPS is effective, particularly if it is from an external validation dataset, if the risk of bias demonstrates inadequate methodology and a high risk of bias, then the premise that the ALPS is effective, is undermined. The surgeon should exercise caution in choosing a leak prediction score.

Studies that develop a model must report its performance either in the development cohort with internal validation or ideally in an external validation set. A model is likely to have good calibration in the development cohort so discrimination is the most (but not the only) important measure of performance. In external validation studies (either within the same paper or in independent study groups testing a model) it is important also to report calibration (agreement between predicted and observed probabilities).[49] This may be achieved with calibration plots in risk deciles with a 'smoothed lowess line' or sometimes in a table.[36 38] The Hosmer-Lemeshow test is a statistical test that can be applied to report level of agreement however it is no longer recommended because it artificially groups patients into categories then generates a p value that has low statistical power and cannot describe the type or extent of miscalibration.[49]

Future research

To optimise existing ALPS, large scale validation studies should be performed, with calibration the key measure. It is neither possible nor realistic to assume that patient populations remain static, so calibration and re-calibration over time and across different geographic populations could optimise performance and result in safer operative decisions.

In model development, it would be prudent to include all patients regardless of diverting stoma placement. Authors should calculate model performance with and without patients with diverting ostomies. This would help to avoid a selection bias by exclusion of an important group of high-risk patients, but address the fact that leaks may be subclinical in the context of stoma diversion so also calculating performance after excluding this group.

Authors of any future ALPS should adhere closely to TRIPOD guidance.[36] This is demanding and highly specified but it would guide authors toward adequate sample sizes, applicability and low risk of bias. Future studies should avoid the routinely missed detail (such as handling of missing data); the methodological pitfalls (such as univariate factor selection before MVLR); the guilt of omission (absent calibration plots) and a mistaken focus (such as on p values for AUC).

The high risk of bias identified in this review is consistent with literature findings of other systematic reviews of prediction models.[50] PROBAST guidelines[10] are clear but have stringent requirements particularly for assessment of analysis bias, and these exacting requirements do not differentiate between a study that achieves eight of nine criteria from one that achieves no criteria; all are considered high risk for bias. Future iterations of PROBAST might consider further dividing the analysis domain into methods of analysis, validation and reporting of model performance.

CONCLUSION

This review provides the reader with an overview of existing ALPS, their strengths and shortcomings. Several models appear to perform well in discriminating patients at highest AL risk but all raise concerns over risk of bias, and nearly all over wider applicability. While we have been able to report the popularity of individual risk factors in ALPS, we are unable to recommend best performing factors because of poor reporting practices and methodological shortcomings. There is potential for effective preoperative and intraoperative risk calculation to guide operative decision-making but selection and re-calibration of the best ALPS with large-scale, precisely reported external validation are needed to benefit colorectal patients.

Twitter Mary L Venn @MaryVenn4 and Dmitri Nepogodiev @dnepo

Contributors DN, CHK, DM and MLV were responsible for the concept of the ALPS study. MLV was the first author of the manuscript and responsible for revisions with the wider study team. The protocol was prepared by MLV who also led data extraction and data analysis together with TP, and CHK advising. MLV and RLH completed statistical analysis. All authors contributed to manuscript review and editing, and approved the final manuscript before submission. MLV is the guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests Conflicts of Interest and Source of Funding: DM is chief investigator of the EAGLE study (ESCP Safe-anastomosis Programme in Colorectal Surgery) that uses a patient risk stratification tool as part of its intervention to reduce anastomotic leak. We thank the European Society of Coloproctology (ESCP) for the overall funding for EAGLE study conduct. RH, CHK, MLV and DN are members of the EAGLE steering and operations committees. This EAGLE study has in no way impacted the selection or assessment of the studies in this review. For the remaining authors no conflicts were declared. This study was unfunded.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Author note Please view this paper's main tables, presented in Supplemental material due to their large size.

ORCID iDs
Mary L Venn http://orcid.org/0000-0002-3015-7039
Richard L Hooper http://orcid.org/0000-0002-1063-0917
Tom Pampiglione http://orcid.org/0000-0003-1156-0401
Dion G Morton http://orcid.org/0000-0001-6784-1689
Dmitri Nepogodiev http://orcid.org/0000-0002-2171-2862
Charles H Knowles http://orcid.org/0000-0001-9854-6754

REFERENCES

1 Manwaring ML, Ko CY, Fleshman JW, *et al*. Identification of consensus-based quality end points for colorectal surgery. *Dis Colon Rectum* 2012;55:294–301.
2 2015 European Society of Coloproctology collaborating group. Relationship between method of anastomosis and anastomotic failure after right Hemicolectomy and Ileo-caecal resection: an international Snapshot audit. *Colorectal Dis* 2017:e296–311.
3 The 2017 European Society of Coloproctology (ESCP) Collaborating Group. The 2017European society of Coloproctology (ESCP) International Snapshot audit of left colon, sigmoid and Rectal Resections - executive summary. *Colorectal Dis* 2018;20:13–4.
4 Glasbey JC, Blanco-Colino R, Kelly M, *et al*. Association of mechanical bowel preparation with oral antibiotics and anastomotic leak following left sided colorectal resection: an international, multicentre, prospective audit. *Colorectal Dis* 2018;20:15–32.
5 Karliczek A, Harlaar NJ, Zeebregts CJ, *et al*. Surgeons lack predictive accuracy for anastomotic leakage in gastrointestinal surgery. *Int J Colorectal Dis* 2009;24:569–76.
6 Sammour T, Lewis M, Thomas ML, *et al*. A simple web-based risk Calculator (Www.Anastomoticleak.Com) is superior to the surgeon's estimate of anastomotic leak after colon cancer resection. *Tech Coloproctol* 2017;21:35–41.

7 Moons KGM, de Groot JAH, Bouwmeester W, *et al*. Critical appraisal and data extraction for systematic reviews of prediction Modelling studies: the CHARMS checklist. *PLoS Med* 2014;11:e1001744.
8 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
9 Souwer ETD, Bastiaannet E, Steyerberg EW, *et al*. Risk prediction models for postoperative outcomes of colorectal cancer surgery in the older population - a systematic review. *J Geriatr Oncol* 2020;11:1217–28.
10 Moons KGM, Wolff RF, Riley RD, *et al*. PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019;170:W1–33.
11 Moons KGM, Wolff RF, Riley RD, *et al*. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1–33.
12 Jiang W, Feng M, Dong X, *et al*. Risk factor analysis on anastomotic leakage after Laparoscopic surgery in Rectal cancer patient with Neoadjuvant therapy and establishment of a Nomogram prediction model] article in Chinese. *Chinese J Gastrointest Surg* 2019;22:748–54.
13 McKenna NP, Bews K, Mathis KL, *et al*. 904 when is diversion indicated after right sided colon resection. *Gastroenterology* 2020;158:1540.
14 Yao HH, Shao F, Huang Q, *et al*. Nomogram to predict anastomotic leakage after Laparoscopic anterior resection with Intracorporeal Rectal Transection and double-stapling technique anastomosis for Rectal cancer. *Hepatogastroenterology* 2014;61:1257–61.
15 Sammour T, Cohen L, Karunatillake AI, *et al*. 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* 2017;21:869–77.
16 Bilimoria KY, Liu Y, Paruch JL, *et al*. Development and evaluation of the universal ACS NSQIP surgical risk Calculator: a decision aid and informed consent tool for patients and Surgeons. *J Am Coll Surg* 2013;217:833–42.
17 Rencuzogullari A, Benlice C, Valente M, *et al*. Predictors of anastomotic leak in elderly patients after Colectomy: Nomogram-based assessment from the American college of Surgeons national surgical quality program procedure-targeted cohort. *Dis Colon Rectum* 2017;60:527–36.
18 McKenna NP, Bews KA, Cima RR, *et al*. Development of a risk score to predict anastomotic leak after left-sided Colectomy: which patients warrant diversion *J Gastrointest Surg* 2020;24:132–43.
19 Crispin A, Klinger C, Rieger A, *et al*. The DGAV risk Calculator: development and validation of statistical models for a web-based instrument predicting complications of colorectal cancer surgery. *Int J Colorectal Dis* 2017;32:1385–97.
20 Yang SU, Park EJ, Baik SH, *et al*. Modified colon leakage score to predict anastomotic leakage in patients who underwent left-sided colorectal surgery. *J Clin Med* 2019;8:1450.
21 Cheng S, He B, Zeng X. Prediction of anastomotic leakage after anterior Rectal resection. *Pak J Med Sci* 2019;35:830–5.
22 Glen S. Ridge regression: simple definition: elementary statistics for the rest of us. 2017. Available: https://www.statisticshowto.com/ridge-regression/
23 Glen S. Lasso regression: simple definition: elementary statistics for the rest of us. 2015. Available: https://www.statisticshowto.com/lasso-regression/
24 Arezzo A, Migliore M, Chiaro P, *et al*. The REAL (Rectal anastomotic leak) score for prediction of anastomotic leak after Rectal cancer surgery. *Tech Coloproctol* 2019;23:649–63.
25 Rojas-Machado SA, Romero-Simó M, Arroyo A, *et al*. Prediction of anastomotic leak in colorectal cancer surgery based on a new Prognostic index PROCOLE (Prognostic colorectal leakage) developed from the meta-analysis of observational studies of risk factors. *Int J Colorectal Dis* 2016;31:197–210.
26 Dekker JWT, Liefers GJ, de Mol van Otterloo JCA, *et al*. Predicting the risk of anastomotic leakage in left-sided colorectal surgery using a colon leakage score. *J Surg Res* 2011;166:e27–34.
27 Soguero-Ruiz C, Hindberg K, Rojo-Alvarez JL, *et al*. Support vector feature selection for early detection of anastomosis leakage from bag-of-words in electronic health records. *IEEE J Biomed Health Inform* 2016;20:1404–15.
28 Soguero-Ruiz C, Hindberg K, Mora-Jiménez I, *et al*. Predicting colorectal surgical complications using heterogeneous clinical data and kernel methods. *J Biomed Inform* 2016;61:87–96.
29 Pasic F, Salkic NN. Predictive score for anastomotic leakage after elective colorectal cancer surgery: A decision making tool for choice of protective measures. *Surg Endosc* 2013;27:3877–82.
30 Penna M, Hompes R, Arnold S, *et al*. Incidence and risk factors for anastomotic failure in 1594 patients treated by Transanal total Mesorectal Excision results from the International TATME Registry. *Ann Surg* 2019;269:700–11.
31 Shen Z, Lin Y, Ye Y, *et al*. The development and validation of a novel model for predicting surgical complications in colorectal cancer of elderly patients: results from 1008 cases. *Eur J Surg Oncol* 2018;44:490–5.
32 Hu X, Cheng Y. A clinical parameters-based model predicts anastomotic leakage after a Laparoscopic total Mesorectal Excision: A large study with data from China. *Medicine* 2015;94:e1003.
33 Yu Z, Liu X, Liu H, *et al*. Impact of pelvic dimensions on anastomotic leak after anterior resection for patients with Rectal cancer. *Surg Endosc* 2021;35:2134–43.
34 Liu X-H, Wu X-R, Zhou C, *et al*. Conversion is a risk factor for postoperative anastomotic leak in Rectal cancer patients - A retrospective cohort study. *Int J Surg* 2018;53:298–303.
35 Xiao C, Zhou M, Yang X, *et al*. Novel Nomogram with Microvascular density in the surgical margins can accurately predict the risk for anastomotic leakage after anterior resection for Rectal cancer. *J Surg Oncol* 2019;120:1412–9.
36 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:735–6.
37 Watanabe T, Miyata H, Konno H, *et al*. Prediction model for complications after low anterior resection based on data from 33,411 japanese patients included in the national clinical database. *Surgery* 2017;161:1597–608. 10.1016/j.surg.2016.12.011 Available: http://dx.doi.org/10.1016/j.surg.2016.12.011
38 van Smeden M, Reitsma JB, Riley RD, *et al*. Clinical prediction models: diagnosis versus prognosis. *J Clin Epidemiol* 2021;132:142–5.
39 Shiwakoti E, Song J, Li J, *et al*. Prediction model for anastomotic leakage after Laparoscopic Rectal cancer resection. *J Int Med Res* 2020;48:0300060520957547.
40 Frasson M, Flor-Lorente B, Rodríguez JLR, *et al*. 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* 2015;262:321–30.
41 Tay K. *What is a Brier score?* 2019. Available: https://statisticaloddsandends.wordpress.com/2019/12/29/what-is-a-brier-score/
42 Zhang Z, Rousson V, Lee W-C, *et al*. Decision curve analysis: a technical NOTE. *Ann Transl Med* 2018;6:308.
43 Park JS, Choi G-S, Kim SH, *et al*. Multicenter analysis of risk factors for anastomotic leakage after Laparoscopic Rectal cancer Excision: the Korean Laparoscopic colorectal surgery study group. *Ann Surg* 2013;257:665–71.
44 Yu X-Q, Zhao B, Zhou W-P, *et al*. Utility of colon leakage score in left-sided colorectal surgery. *J Surg Res* 2016;202:398–402.
45 Muñoz NP, Rodríguez MG, Pérez-Castilla A, *et al*. Applicability of the colon leakage score as an anastomotic leakage Predictor in colorectal cancer surgery. *Rev Chil Cir* 2018;70:439–44.
46 Hoshino N, Hida K, Sakai Y, *et al*. Nomogram for predicting anastomotic leakage after low anterior resection for Rectal cancer. *Int J Colorectal Dis* 2018;33:411–8.
47 Zheng H, Wu Z, Wu Y, *et al*. Laparoscopic surgery may decrease the risk of clinical anastomotic leakage and a Nomogram to predict anastomotic leakage after anterior resection for Rectal cancer. *Int J Colorectal Dis* 2019;34:319–28.
48 Kalkum E, Klotz R, Seide S, *et al*. Systematic reviews in surgery—recommendations from the study center of the German society of surgery. *Langenbecks Arch Surg* 2021;406:1723–31.
49 Van Calster B, McLernon DJ, van Smeden M, *et al*. Calibration: the Achilles heel of predictive Analytics. *BMC Med* 2019;17:230:230.:.
50 Wynants L, Van Calster B, Collins GS, *et al*. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. *BMJ* 2020;369:m1328.
51 Liu Y, Wan X, Wang G, *et al*. A scoring system to predict the risk of anastomotic leakage after anterior resection for Rectal cancer. *J Surg Oncol* 2014;109:122–5.
52 Kim CH, Lee SY, Kim HR, *et al*. Nomogram prediction of anastomotic leakage and determination of an effective surgical strategy for reducing anastomotic leakage after Laparoscopic Rectal cancer surgery. *Gastroenterol Res Pract* 2017;2017:4510561.
53 Klose J, Tarantino I, von Fournier A, *et al*. A Nomogram to predict anastomotic leakage in open Rectal surgery—hope or Hype *J Gastrointest Surg* 2018;22:1619–30.