European Journal of Surgical Oncology 50 (2024) 108338

Contents lists available at ScienceDirect

European Journal of Surgical Oncology

journal homepage: www.ejso.com

Development and external validation of preoperative clinical prediction models for postoperative outcomes including preoperative aerobic fitness in patients approaching elective colorectal cancer surgery

Anne C.M. Cuijpers [a,b], Tim Lubbers [a,b], Jaap J. Dronkers [c], Aniek F.J.M. Heldens [d], Siebrand B. Zoethout [e], Duncan Leistra [f], Sander M.J. van Kuijk [g], Nico L.U. van Meeteren [h,i], Laurents P.S. Stassen [a,j], Bart C. Bongers [j,k,*]

[a] Department of Surgery, Maastricht University Medical Centre, Maastricht, the Netherlands
[b] Department of Surgery, GROW, Research Institute for Oncology and Reproduction, Maastricht University, Maastricht, the Netherlands
[c] Expertise Centre Healthy Urban Living, Research Group Innovation of Human Movement Care, HU University of Applied Sciences Utrecht, Utrecht, the Netherlands
[d] Department of Physical Therapy, Maastricht University Medical Centre, Maastricht, the Netherlands
[e] Department of Physical Therapy, Deventer Hospital, Deventer, the Netherlands
[f] Department of Physical Therapy, Nij Smellinghe Hospital, Drachten, the Netherlands
[g] Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Centre, Maastricht, the Netherlands
[h] Top Sector Life Sciences and Health (Health~Holland), The Hague, the Netherlands
[i] Department of Anesthesiology, Erasmus Medical Centre, Rotterdam, the Netherlands
[j] Department of Surgery, NUTRIM, Institute of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, the Netherlands
[k] Department of Nutrition and Movement Sciences, NUTRIM, Institute of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, the Netherlands

ARTICLE INFO

Keywords:
Colorectal cancer
Aerobic fitness
Complication risk
Recovery of physical functioning
Prediction model
Prehabilitation

ABSTRACT

Introduction: Preoperative aerobic fitness is associated with postoperative outcomes after elective colorectal cancer (CRC) surgery. This study aimed to develop and externally validate two clinical prediction models incorporating a practical test to assess preoperative aerobic fitness to distinguish between patients with and without an increased risk for 1) postoperative complications and 2) a prolonged time to in-hospital recovery of physical functioning after elective colorectal cancer (CRC) surgery.
Materials and methods: Models were developed using prospective data from 256 patients and externally validated using prospective data of 291 patients. Postoperative complications were classified according to Clavien-Dindo. The modified Iowa level of assistance scale (mILAS) was used to determine time to postoperative in-hospital physical recovery. Aerobic fitness, age, sex, body mass index, American Society of Anesthesiologists (ASA) classification, neoadjuvant treatment, surgical approach, tumour location, and preoperative haemoglobin level were potential predictors. Areas under the curve (AUC), calibration plots, and Hosmer-Lemeshow tests evaluated predictive performance.
Results: Aerobic fitness, sex, age, ASA, tumour location, and surgical approach were included in the final models. External validation of the model for complications and postoperative recovery presented moderate to fair discrimination (AUC 0.666 (0.598–0.733) and 0.722 (0.651–0.794), respectively) and good calibration. High sensitivity and high negative predictive values were observed in the lower predicted risk categories (<40 %).
Conclusion: Both models identify patients with and without an increased risk of complications or a prolonged time to in-hospital physical recovery. They might be used for improving patient-tailored preoperative risk assessment and targeted and cost-effective application of prehabilitation interventions.

* Corresponding author. Department of Nutrition and Movement Sciences, Department of Surgery Maastricht University PO Box 616, 6200 MD Maastricht, the Netherlands.
E-mail addresses: a.cuijpers@maastrichtuniversity.nl (A.C.M. Cuijpers), tim.lubbers@mumc.nl (T. Lubbers), jaap.dronkers@hu.nl (J.J. Dronkers), aniek.heldens@mumc.nl (A.F.J.M. Heldens), s.zoethout@dz.nl (S.B. Zoethout), d.leistra@nijsmellinghe.nl (D. Leistra), sander.van.kuijk@mumc.nl (S.M.J. van Kuijk), meeteren@health-holland.com (N.L.U. van Meeteren), lps.stassen@mumc.nl (L.P.S. Stassen), bart.bongers@maastrichtuniversity.nl (B.C. Bongers).

https://doi.org/10.1016/j.ejso.2024.108338
Received 19 December 2023; Received in revised form 1 March 2024; Accepted 9 April 2024
Available online 11 April 2024

1. Introduction

Colorectal cancer (CRC) is highly prevalent and is predominantly cured by surgical resection [1,2]. The pathophysiological processes of CRC and the subsequent surgical procedure provide a substantial level of stress to the body. Especially high-risk patients, based on a low preoperative physical fitness, high comorbidity burden, and/or high age, are prone to perioperative functional decline and postoperative complications [3,4]. With an aging general population, and consequently, an increasing number of potentially high-risk patients undergoing elective abdominal surgery, it is essential to reduce treatment-related complications and maintain vital functioning, including physical performance, participation in daily activities, and quality of life before and after surgery [5,6].

Indicators of preoperative physical fitness, especially aerobic fitness as assessed using cardiopulmonary exercise testing (CPET), have consistently been proven to be associated with morbidity and recovery after CRC surgery and may therefore contribute to improved preoperative risk assessment [7–10]. However, preoperative CPET is not part of usual care in the majority of hospitals and is relatively expensive in terms of required equipment and trained personnel. The use of measurement tools that are easier to perform and implement will help identify patients with low aerobic fitness more easily. As a result, interventions to optimize or maintain preoperative physical (aerobic) fitness and reduce postoperative morbidity risk can be offered faster and to more patients [11–13]. The steep ramp test (SRT) is a short-time maximal exercise test correlated with aerobic fitness as assessed by CPET [14] and associated with postoperative outcomes in patients with CRC [10]. It seems an accurate and more practical test to evaluate cardiorespiratory fitness and to more easily predict the risk of postoperative morbidity in patients scheduled for abdominal surgery [10,14,15].

This study aimed to develop and externally validate two clinical prediction models combining SRT performance with other patient-related risk factors to distinguish between patients with and without a preoperatively increased risk for 1) postoperative complications, and 2) a prolonged in-hospital physical recovery time after elective CRC surgery.

2. Methods

This multicentre prospective observational cohort study (PROCLINA) was conducted at the Maastricht University Medical Centre+, Gelderse Vallei hospital, Nij Smellinghe hospital, and Deventer hospital, all in The Netherlands. The study was approved by the Medical Ethical Committee of Maastricht University Medical Centre/Maastricht University (15-4-234). The study is reported using the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement [16,17]. Data were collected from January 2016 until March 2020. Data from the Maastricht University Medical Centre were previously used to evaluate the associations between preoperative aerobic fitness and postoperative outcomes in patients scheduled for elective CRC surgery [10]; for the current study, these data were used as derivation cohort to develop the logistic regression model. Data collected at the other three participating hospitals was used as cohort for external validation.

2.1. Study population

As part of usual care, all consecutive patients diagnosed with CRC who were scheduled for elective tumour resection in one of the participating hospitals were referred to the outpatient clinic of the physical therapy department for a preoperative physical fitness assessment. The inclusion and exclusion criteria for this study are outlined in our initial report [10]. Data of patients aged ≥18 years who gave informed consent for using their usual care data for research purposes were eligible for inclusion in the study. Exclusion criteria were preoperative assessment of physical fitness before start of neoadjuvant chemo-radiotherapy or >2 months prior to surgery, physical exercise training before surgery (exercise prehabilitation), total pelvic exenteration, no bowel resection due to peritoneal metastases, or postoperative air-fluidized sand bed therapy.

2.2. Perioperative data collection

Except from the preoperative physical fitness assessment, CRC care is offered similarly throughout the Netherlands in accordance to standardized national guidelines and the enhanced recovery after surgery (ERAS) protocol [18]. Perioperative data including patient characteristics, tumour characteristics, treatment details, and postoperative outcomes are prospectively recorded in the nationwide Dutch ColoRectal Audit (DCRA) for benchmark purposes (Dutch Institute for Clinical Auditing, Leiden, the Netherlands). Data related to preoperative physical fitness as measured during the preoperative physical fitness assessment and data on postoperative in-hospital physical recovery were registered in the PROCLINA database by MRDM (Medical Data Research Management B.V., Deventer, the Netherlands), in which the data importer was blinded for the patient's DCRA data. After completion of data collection, the PROCLINA and DCRA databases were combined on an individual patient level by the data processor (MRDM) and handed back to the researchers to check for correctness and completeness. Where possible, missing values were added based on data from the electronic patient files.

2.3. Preoperative aerobic fitness assessment

As part of a preoperative physical fitness assessment, aerobic fitness was measured using the modified SRT, a short-time maximal exercise test on a calibrated cycle ergometer (Lode Corival Rehab, Lode BV, Groningen, the Netherlands). Patients were instructed to pedal between 70 and 80 revolutions/min while, after a 2-min warm-up of unloaded cycling, the work rate was increased with increments of 10 W/10 s in a ramp like manner of 1 W/s [10,15,19,20]. The test continued until the pedalling frequency definitely dropped below 60 revolutions/min despite strong verbal encouragement. SRT performance, expressed as the attained peak work rate adjusted for body mass (WR_{peak} in W/kg) was used as measure for preoperative aerobic fitness.

2.4. Postoperative outcomes

Thirty-day postoperative complications were classified using the Clavien-Dindo (CD) classification [21]. Starting from postoperative day one, all patients received postoperative guidance by a physical therapist, on average 20 min per day. In-hospital physical recovery was monitored daily using the modified Iowa level of assistance scale (mILAS) [22]. The mILAS measures the ability to perform five daily activities necessary for independent physical functioning (supine-to-sit, sit-to-supine, sit-to--stand, walking, and stair-climbing; the latter only when essential in domestic circumstances). The level of assistance needed for all transfers were scored from 0 (independent) to 6 (not tested). The sum of the five individual scores delivers the mILAS score, which ranges from 0 to 30. Higher scores indicate more assistance, whereas a score of 0 indicates a patient is functionally recovered and ready for discharge home from a physical therapy point of view. Postoperative time to in-hospital physical recovery was defined as the time in days between the day of surgery and the day a patient reached a mILAS score of 0 (time to mILAS = 0).

2.5. Sample size calculation

As a rule of thumb for development of the logistic regression models, the 10 events per predictor variable rule was used [23]. Development of any postoperative complication (CD ≥ I) and a time to mILAS = 0 higher than the median value of the sample were defined as events. Based on clinical relevance, expert opinion, and previous literature, SRT performance (WR_{peak} in W/kg), age (years), sex (male/female), body mass index (BMI), American Society of Anesthesiologists (ASA) classification, neoadjuvant treatment (radiotherapy and/or chemotherapy), planned surgical approach (laparoscopy/robot or laparotomy), tumour location (colon or rectum), and preoperative haemoglobin level (g/dl) were identified as potential preoperative predictors for postoperative complications or prolonged time to in-hospital physical recovery [10,24]. Assuming a 40 % event rate for postoperative complications [10] and evaluating at least 9 potential predictors, 225 patients would be required to achieve stable estimates.

2.6. Statistical analysis

Statistical analysis, model development, and external validation was performed using IBM SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, NY, USA). Missing values were imputed using stochastic regression imputation with fully conditional specification. Normality of continuous variables was tested using histograms, Q-Q plots, and Kolmogorov-Smirnov tests. Values were displayed as mean ± standard deviation (SD) or median and interquartile range [IQR]. According to normality, the independent-samples *t*-test or the Mann-Whitney *U* test was used to test for between-group differences in case of continuous variables. Categorical values were displayed as number (percentage) and between-group differences were analysed using Pearson's $\chi 2$ test. For prediction modelling purposes, postoperative outcomes were dichotomized. Postoperative complications were classified as no complication (CD = 0) or any complication (CD ≥ I). Postoperative time to in-hospital physical recovery was dichotomized based on median time to mILAS = 0 (in days) in the derivation cohort. A time to mILAS = 0 higher than the median value of the sample was defined as prolonged in-hospital physical recovery time. Prediction models were developed using multivariable logistic regression analysis. Stepwise backward elimination was used to eliminate non-significant predicting variables from the model. A liberal p-value of 0.200 was used to prevent premature elimination of potential predictors [25]. Discriminating ability of the prediction model was assessed by the area under the receiver operating characteristic curve (AUC). Calibration of the model was assessed using the Hosmer and Lemeshow test and calibration plots for goodness of fit. The prediction model was then applied to the validation set and a new assessment of discrimination and calibration was calculated to evaluate the predictive performance.

3. Results

3.1. Population characteristics

The derivation cohort consisted of 256 patients who were eligible for analysis (Fig. 1), of which 107 (41.8 %) patients developed any postoperative complication. Median time between physical fitness assessment and surgery was 11 days [6.00, 17.00]. Median postoperative time to in-hospital physical recovery was 4 days [3.00; 7.75]. In total, 3.0 % of the values were missing in the derivation cohort. A total of 82 patients (32.0 %) in the derivation cohort had one or more missing values. The validation cohort consisted of 291 eligible patients. In this cohort, 2.9 % of the values in 85 patients (29.2 %) were missing. The validation cohort differed statistically significant from the derivation cohort in BMI, ASA classification, preoperative haemoglobin levels, neoadjuvant treatment, and tumour location. A total of 92 patients (31.6 %) in the validation cohort developed postoperative complications. Median postoperative time to in-hospital physical recovery was 3 days [2.00; 4.00]. Baseline characteristics of both cohorts are presented in Table 1.

3.2. Development of the prediction models

After stepwise backward elimination in the derivation cohort, five independent predictors were retained in the final prediction model for the development of postoperative complications: sex, ASA classification, SRT performance, tumour location and planned surgical approach. Patients had a higher probability of developing postoperative complications based on sex (odds ratio (OR) 0.497, 95 % confidence interval (CI) 0.281–0.878 for female sex), a higher ASA classification (OR 1.857, 95 % CI 1.097–3.143), a lower preoperative aerobic fitness (OR 0.646, 95 %

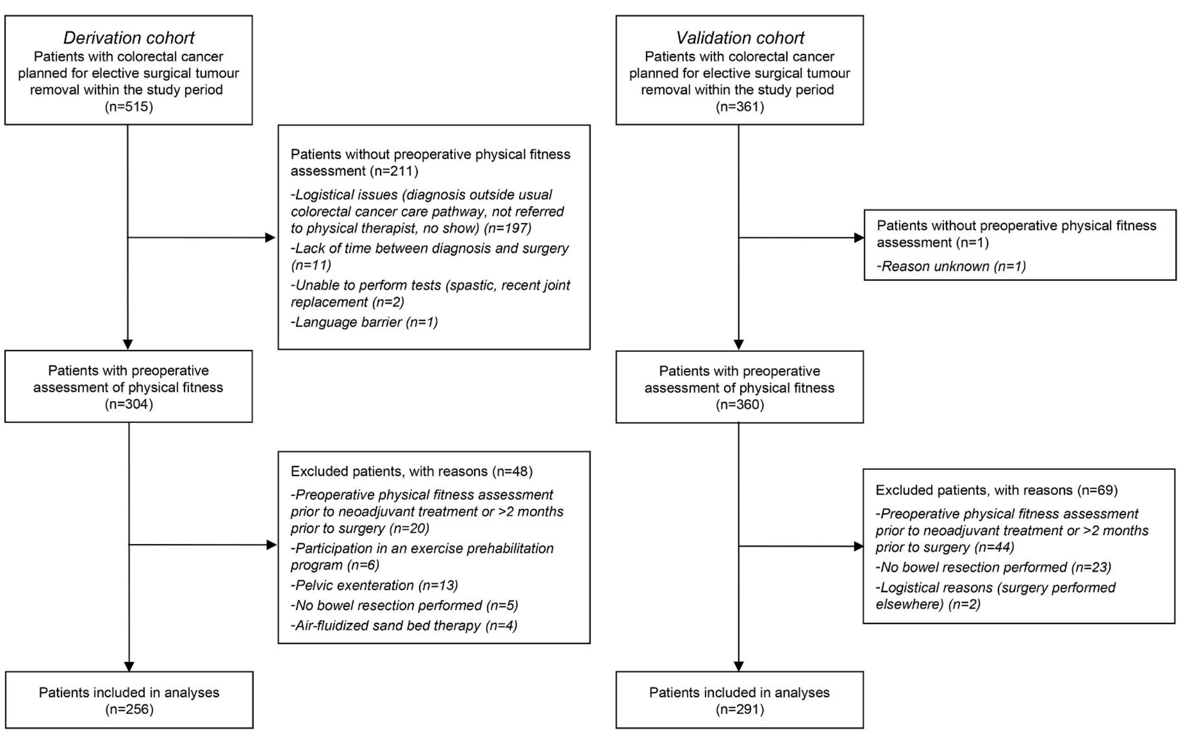

Fig. 1. Inclusion flow derivation and validation cohort.

Table 1
Preoperative baseline characteristics and postoperative outcomes in the derivation and validation cohort and differences between the derivation cohort and external validation cohort.

Variable	Derivation cohort (n = 256)		Validation cohort (n = 291)		p-value
Age (years)	69.4	(±10.0)	68.6	(±8.9)	0.329
Sex					0.616
Male	145	(56.6 %)	171	(58.8 %)	
Female	111	(43.4 %)	120	(41.2 %)	
BMI (kg/m^2)	26.9	(±5.0)	27.9	(±4.7)	***0.023***
Hb (g/dl)	12.8	(±2.0)	13.4	(±1.9)	***0.002***
ASA classification					***0.034***
I	23	(9.0 %)	42	(14.4 %)	
II	164	(64.1 %)	174	(59.8 %)	
III	69	(27.0 %)	70	(24.1 %)	
IV	0	(0 %)	5	(1.7 %)	
Neoadjuvant therapy	62	(24.2 %)	7	(2.4 %)	***<0.001***
Tumour location					***<0.001***
Colon	165	(64.5 %)	233	(80.1 %)	
Rectum	91	(35.5 %)	58	(19.9 %)	
Planned surgical approach					0.130
Laparoscopy/robot (assisted)	230	(89.8 %)	249	(85.6 %)	
Laparotomy	26	(10.2 %)	42	(14.4 %)	
Preoperative aerobic fitness					0.679
SRT WR_{peak} (W/kg)	2.137	(±0.800)	2.110	(±0.774)	
Postoperative outcomes					
Conversion	34[b]	(14.8 %)	8[c]	(3.2 %)	***<0.001***
Complications (CD ≥ I) within 30 days	107	(41.8 %)	92	(31.6 %)	***0.014***
CD classification					***0.003***
I	21	(8.2 %)	27	(9.3 %)	
II	42	(16.4 %)	38	(13.1 %)	
IIIa	9	(3.5 %)	6	(2.1 %)	
IIIb	19	(7.4 %)	5	(1.7 %)	
IVa	12	(4.7 %)	5	(1.7 %)	
IVb	2	(0.8 %)	8	(2.7 %)	
V	2	(0.8 %)	3	(1.0 %)	
Readmission rate within 90 days	31	(12.1 %)	17	(5.8 %)	***0.010***
Time to mILAS = 0 (days)	4.0	[3.0; 7.75]	3.0	[2.0; 4.0]	***<0.001***[a]
Time to mILAS = 0 > 4 days	111	(43.4 %)	56	(19.2 %)	***<0.001***
Length of hospital stay (days)	6.0	[4.0; 11.0]	5.0	[4.0; 7.0]	***0.001***[a]

Data displayed as mean (±standard deviation), median [interquartile range], or number (percentage).
Abbreviations: ASA: American Society of Anesthesiologists classification, BMI: body mass index, CD: Clavien-Dindo, Hb: haemoglobin, mILAS: modified Iowa level of assistance scale, SRT: steep ramp test, WR_{peak}: work rate at peak exercise.
[a] Mann-Whitney *U* test.
[b] n = 230.
[c] n = 249.

CI 0.436–0.956 for each 1 W/kg increase in SRT WR_{peak}), a tumour located in the rectum (OR 2.578, 95 % CI 1.467–4.531), and surgery performed via laparotomy (OR 2.063, 95 % CI 0.853–4.992) (Table 2). The model had an AUC of 0.713 (95 % CI 0.649–0.778). The discriminative abilities and calibration of the model are presented in Fig. 2A and C. The calibration plot closely follows the 45-degree line indicating good calibration. In addition, the Hosmer-Lemeshow goodness-of-fit test was non-significant ($p = 0.158$), indicating no evidence of deviation of good model fit, with a Nagelkerke R^2 of 0.156.

The prediction model for a prolonged in-hospital physical recovery time consisted of five independent predictors after stepwise backwards elimination: age, ASA classification, SRT performance, tumour location, and planned surgical approach. Prolonged in-hospital physical recovery was defined as >4 days to reach a mILAS score of 0. Patients had a higher probability of a prolonged postoperative in-hospital physical recovery based on age (OR 0.974, 95 % CI 0.943–1.006 for each 1-year increase in age), a higher ASA classification (OR 2.003, 95 % CI 1.164–3.448), a lower preoperative aerobic fitness (OR 0.475, 95 % CI 0.307–0.736 for each 1 W/kg increase in SRT WR_{peak}), a tumour located in the rectum (OR 2.793, 95 % CI 1.560–5.002) and surgery performed via laparotomy (OR 3.482, 95 % CI 1.337–9.069) (Table 2). The model had an AUC 0.735 (0.674–0.796; $p < 0.001$). The discriminative abilities and calibration are presented in Fig. 3A and C. The calibration plot closely follows the 45-degree line indicating good calibration and the Hosmer-Lemeshow goodness-of-fit test was non-significant ($p = 0.453$), with a Nagelkerke R^2 of 0.216.

Table 2
Preoperative predictors of A: postoperative complications (CD ≥ I), and B: a prolonged postoperative time to in-hospital recovery of physical functioning (time to mILAS = 0 > 4 days) using logistic regression analysis.

A: postoperative complications (CD ≥ I)	B	OR	95 % CI	p-value
Total model[a]				
Age	−0.019	0.981	0.949–1.015	0.278
Sex	−0.775	0.461	0.249–0.851	0.013
BMI (kg/m^2)	−0.027	0.974	0.912–1.040	0.425
ASA classification	0.662	1.938	1.116–3.366	0.019
SRT WR_{peak} (W/kg)	−0.689	0.502	0.290–0.870	0.014
Hb (g/dl)	0.122	1.129	0.892–1.430	0.312
Tumour location	0.752	2.122	0.990–4.547	0.053
Surgical approach	0.844	2.325	0.934–5.791	0.070
Neoadjuvant therapy	0.227	1.255	0.546–2.886	0.593
Intercept	0.632			
Backward stepwise (p <0.200)[a]				
Sex	−0.700	0.497	0.281–0.878	0.016
ASA classification	0.619	1.857	1.097–3.143	0.021
SRT WR_{peak} (W/kg)	−0.438	0.646	0.436–0.956	0.029
Tumour location	0.947	2.578	1.467–4.531	0.001
Surgical approach	0.724	2.063	0.853–4.992	0.108
Intercept	−0.893			

B: prolonged time to in-hospital physical recovery (mILAS = 0 > 4 days)	B	OR	95 % CI	p-value
Total model[a]				
Age	−0.028	0.972	0.939–1.006	0.108
Sex	−0.227	0.797	0.432–1.469	0.467
BMI (kg/m^2)	0.016	1.016	0.949–1.088	0.640
ASA classification	0.696	2.006	1.140–3.532	0.016
SRT WR_{peak} (W/kg)	−0.738	0.478	0.271–0.843	0.011
Hb (g/dl)	−0.021	0.979	0.844–1.137	0.782
Tumour location	1.239	3.452	1.545–7.713	0.003
Surgical approach	1.303	3.679	1.377–9.829	0.009
Neoadjuvant therapy	−0.305	0.737	0.311–1.748	0.489
Intercept	1.149			
Backward stepwise (p <0.200)[a]				
Age	−0.027	0.974	0.943–1.006	0.104
ASA classification	0.695	2.003	1.164–3.448	0.012
SRT WR_{peak} (W/kg)	−0.744	0.475	0.307–0.736	0.001
Tumour location	1.027	2.793	1.560–5.002	0.001
Surgical approach	1.248	3.482	1.337–9.069	0.011
Intercept	1.131			

Abbreviations: ASA: American Society of Anesthesiologists classification, BMI: body mass index, CI: confidence interval, CD: Clavien-Dindo, Hb: haemoglobin, mILAS: modified Iowa level of assistance scale, OR: odds ratio, SRT: steep ramp test, WR_{peak}: work rate at peak exercise.
[a] Age is expressed in years; for sex: male = 0, female = 1; for ASA classification: 1 = 1, 2 = 2, 3 = 3, 4 = 4; SRT WR_{peak} is expressed in W/kg; for tumour location: colon = 0, rectum = 1; for surgical approach: laparoscopy or robot (assisted) = 0, laparotomy = 1.

3.3. External validation of the prediction models

Validation of the prediction model for development of postoperative complications showed an AUC of 0.666 (0.598–0.733; $p < 0.001$) in the independent validation cohort (Fig. 2B). For the prediction model for

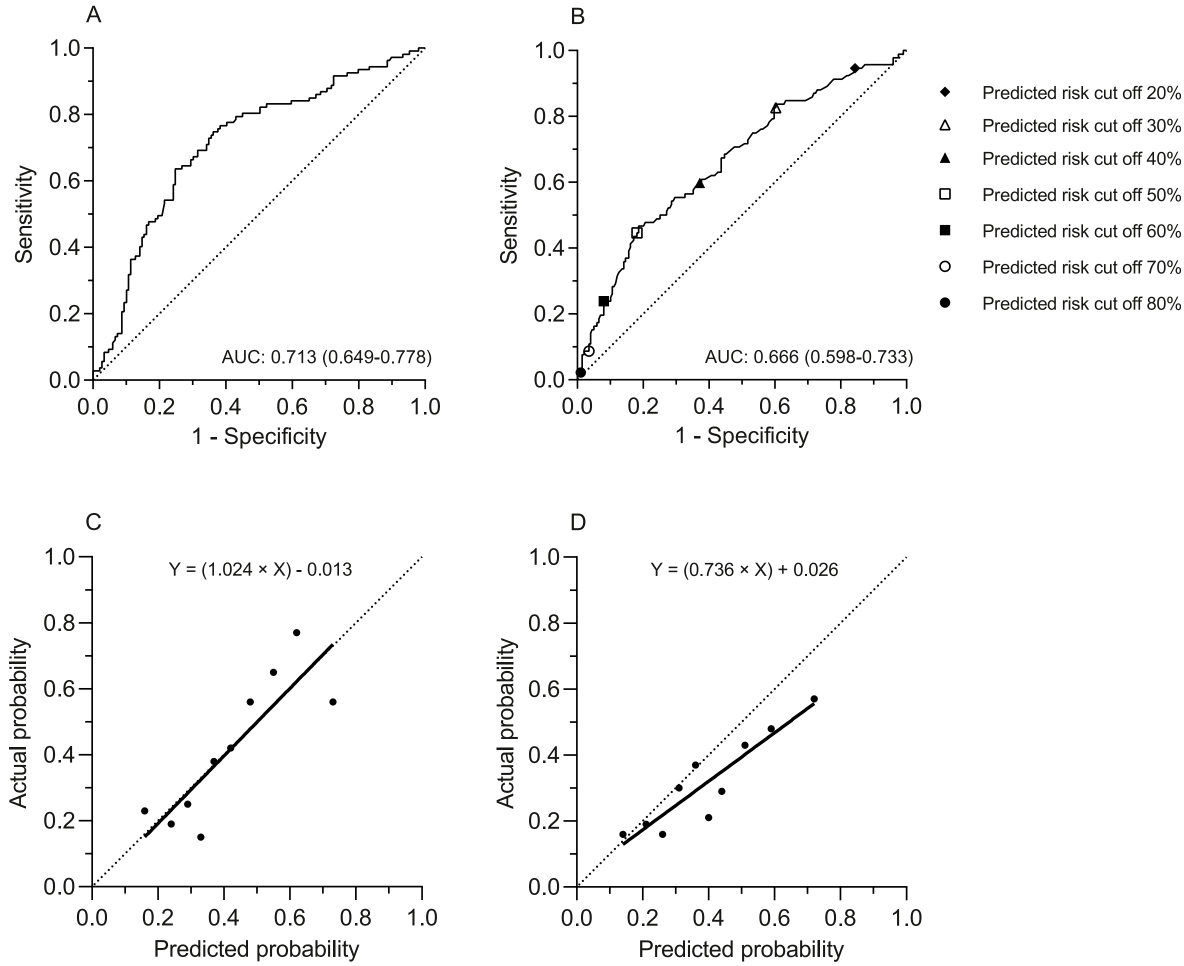

Fig. 2. Discriminative and calibration abilities of prediction model for postoperative complications: ROC curve derivation cohort (graph A); ROC curve validation cohort including multiple risk thresholds (graph B); calibration plot derivation cohort (graph C); calibration plot validation cohort (graph D).

prolonged postoperative in-hospital physical recovery, an AUC of 0.722 (0.651–0.794; $p < 0.001$) was achieved (Fig. 3B). Calibration curves of both models showed, on average, a slight overestimation of the risk (Figs. 2D and 3D, respectively).

3.4. Clinical usefulness

Table 3 shows the model formulas and model characteristics for several threshold values of predictive probabilities to distinguish patients at low and high risk for postoperative complications and a prolonged time to in-hospital physical recovery. Threshold values are projected on the receiver operating curves (ROC) and displayed in Figs. 2B and 3B, respectively. Sensitivity, specificity, and the positive and negative predictive value were estimated for different risk thresholds. At the lower thresholds (>20, >30, and >40 %), high sensitivity and high negative predictive values were observed, suggesting the ability of the models to identify patients at low risk for developing postoperative complications, or experiencing a prolonged postoperative in-hospital time to physical recovery. Except for SRT performance, the predictors in both prediction models are not modifiable. As shown in Fig. 4, SRT performance is inversely related to the predicted probabilities of both prediction models in both the derivation and calibration cohort.

4. Discussion

This study developed and externally validated two models to preoperatively distinguish between patients with and without 1) an increased risk for postoperative complications and 2) a prolonged time to in-hospital physical recovery after elective CRC surgery. Both models demonstrated fair to moderate discriminatory abilities in the external validation cohort with AUCs of 0.666 and 0.722, respectively. Calibration of both models was slightly overfitted because of an overestimation of the predicted risk in the validation cohort.

Both models showed high sensitivity and high negative predictive values in the low-risk threshold categories. This indicated that both models are useful to preoperatively identify patients with a low risk of postoperative complications or a low risk a prolonged postoperative time to in-hospital recovery of physical functioning. In both models, preoperative aerobic fitness appeared to be an important risk predictor. Unlike the other relevant predictors, preoperative aerobic fitness is a modifiable risk predictor. The strong inverse relationship between the level of preoperative aerobic fitness and the predicted risk for developing complications, as well as for experiencing a prolonged postoperative time to in-hospital physical recovery suggests that improving preoperative aerobic fitness can reduce the risk for postoperative morbidity and accelerate postoperative in-hospital recovery of physical functioning. A high sensitivity and high negative predictive value are preferable characteristics of a prediction model to select patients who not necessarily need prehabilitation and might go straight into surgery. Based on these model characteristics, only a very small number of patients will receive a low predicted score while being at risk for unfavourable postoperative outcomes. From a cost-effectiveness perspective, it can be argued that also a high specificity is preferable. However, from

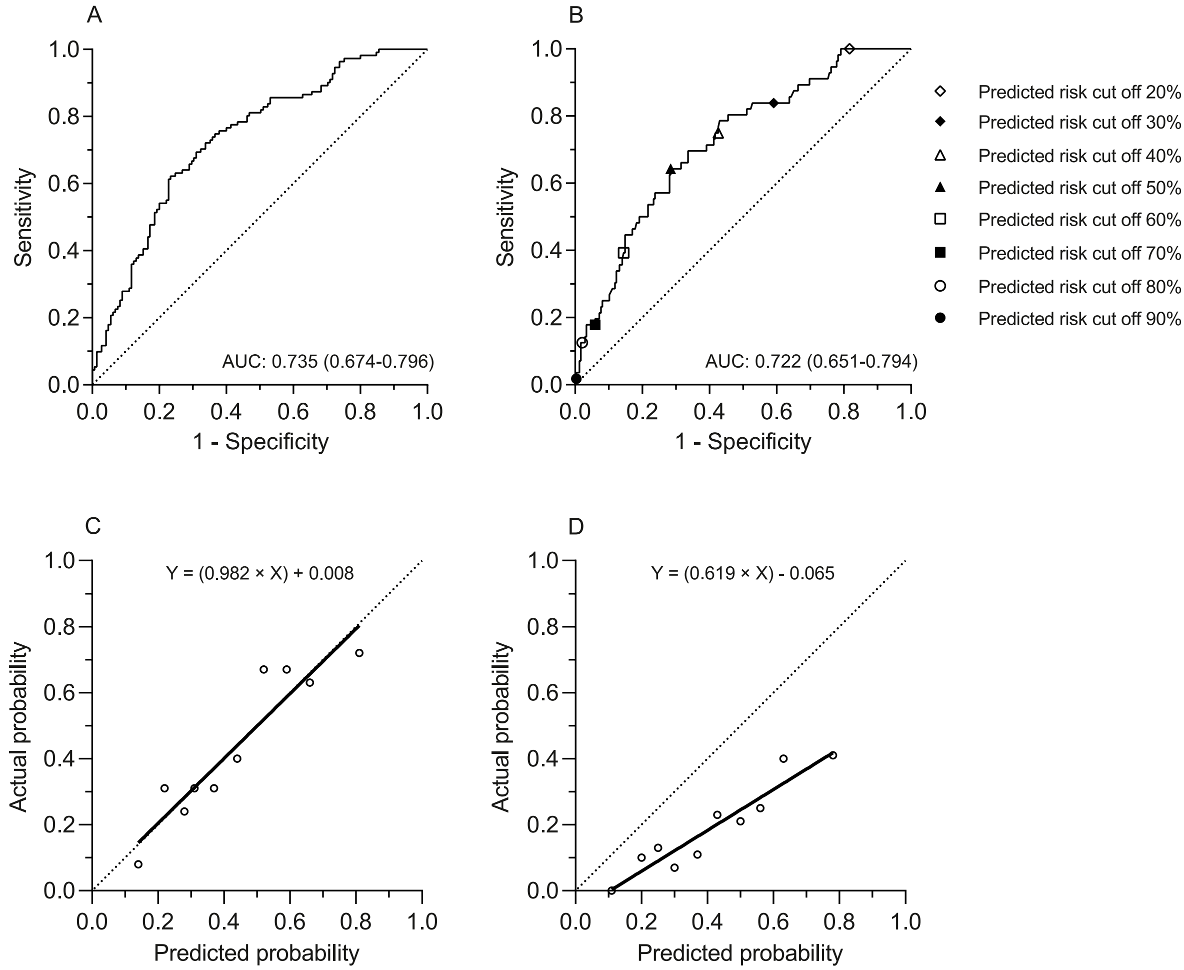

Fig. 3. Discriminative and calibration abilities of prediction model for postoperative in-hospital physical recovery: ROC curve derivation cohort (graph A); ROC curve validation cohort including multiple risk thresholds (graph B); calibration plot derivation cohort (graph C); calibration plot validation cohort (graph D).

a clinical point of view, false positive cases are less of a concern when selecting patients who might benefit from prehabilitation, because the consequence is that some low-risk patients participate in a prehabilitation program.

Good preoperative aerobic fitness is associated with a reduced risk for postoperative morbidity after major abdominal surgery [9,10,26]. However, current literature regarding an externally validated prediction of postoperative morbidity based on preoperative physical (aerobic) fitness and selecting patients who might benefit from exercise prehabilitation is sparse. Previous literature identified CPET, which is the gold standard for aerobic fitness assessment, as an accurate risk predictor for postoperative complications after CRC surgery [8]. Especially oxygen uptake (VO_2) at the ventilatory anaerobic threshold and VO_2 at peak exercise (VO_{2peak}) as measured by CPET before major abdominal surgery appeared to have good discriminative abilities in predicting postoperative complications [8,9,26–30]. External validation of the predictive abilities of these CPET-derived variables in CRC patients revealed an AUC of 0.79 (95 % CI 0.76–0.83; $p < 0.001$) and 0.77 (95 % CI 0.72–0.82; $p < 0.001$) for VO_2 at the ventilatory anaerobic threshold and VO_{2peak}, respectively [8]. The incremental shuttle walk test, a maximal field exercise test, has also been reported as a potentially useful predictor for postoperative complications (AUC 0.755, 95 % CI 0.592–0.918; $p = 0.027$), but was only measured in small patient populations without external validation of the predicting abilities [7,31]. Compared to the current model, the predictive abilities of CPET-derived variables seem better. However, CPET is often not feasible to perform in all institutions and might therefore not be preferred as a standard method of assessment in all hospitals. Compared to CPET, the SRT is a short and easily accessible maximal exercise test to estimate preoperative aerobic fitness and is equally useful to prescribe training intensity and objectively measure training progress [14,15,32]. Therefore, the current prediction model including SRT performance might be an externally valid alternative and more practical tool to distinguish patients at low risk from patients at a high risk for postoperative morbidity. Moreover, the SRT can be used to guide short-term high-intensity interval training to preoperatively improve the cardiorespiratory fitness of individual patients in routine practice to improve patient-related and treatment-related outcomes [14].

To our knowledge, no previous study presented externally validated data to preoperatively predict postoperative in-hospital physical recovery time after major abdominal surgery. The presented model preoperatively distinguished between patients at low risk for a prolonged time to in-hospital physical recovery (time to a mILAS score of 0 > 4 days). This was based on preoperative morbidity-, surgery-, and tumour-related predictors. Also, it is found to be independent of potential postoperative complications that might also prolong postoperative recovery. In previous literature, it has been shown that patients with a higher preoperative physical fitness are more likely to experience a faster postoperative recovery [10,30,33]. As preoperative aerobic fitness is the only modifiable risk factor in both prediction models developed in the current study, improving preoperative aerobic fitness might improve postoperative outcomes in high-risk patients. One could question the additional value of a separate prediction model for time to postoperative clinical recovery of physical functioning, as the risk predictors in both

Table 3
Model characteristics for multiple risk thresholds.

Threshold	Validation cohort (n = 291)			
Predicted postoperative complication risk	Sensitivity	Specificity	PPV	NPV
>20 %	94.6 %	15.6 %	34.1 %	86.1 %
>30 %	82.6 %	39.7 %	38.8 %	83.2 %
>40 %	59.8 %	62.8 %	42.6 %	77.2 %
>50 %	44.6 %	81.9 %	53.2 %	76.2 %
>60 %	23.9 %	92.0 %	57.9 %	76.2 %
>70 %	8.7 %	96.5 %	53.3 %	69.6 %
>80 %	2.2 %	99.0 %	50.0 %	68.6 %

Formula: P(postoperative complications) = 1/(1 + exp(-(-0.893 - (0.700 × sex) + (0.691 × ASA classification) - (0.438 × SRT WR_{peak}) + (0.947 × tumour location) + (0.724 × surgical approach))))

For sex: male = 0, female = 1; for ASA classification: 1 = 1, 2 = 2, 3 = 3, 4 = 4; SRT WR_{peak} is expressed in W/kg; for tumour location: colon = 0, rectum = 1; for surgical approach: laparoscopy or robot (assisted) = 0, laparotomy = 1.

Predicted postoperative risk time to mILAS = 0 > 4 days	Sensitivity	Specificity	PPV	NPV
>20 %	100.0 %	18.3 %	22.7 %	100.0 %
>30 %	83.9 %	40.9 %	25.3 %	91.4 %
>40 %	75.0 %	57.4 %	29.6 %	90.6 %
>50 %	64.3 %	71.5 %	35.0 %	89.4 %
>60 %	39.3 %	85.5 %	39.3 %	85.5 %
>70 %	17.9 %	94.0 %	41.7 %	82.8 %
>80 %	12.5 %	97.9 %	58.3 %	82.4 %
>90 %	1.8 %	99.6 %	50.0 %	81.0 %

Formula: P(time to mILAS = 0 > 4 days) = 1/(1 + exp(-(1.131 - (0.027 × age) + (0.695 × ASA classification) - (0.744 × SRT WR_{peak}) + (1.027 × tumour location) + (1.248 × surgical approach))))

Age is expressed in years; for ASA classification: 1 = 1, 2 = 2, 3 = 3, 4 = 4; SRT WR_{peak} is expressed in W/kg; for tumour location: colon = 0, rectum = 1; for surgical approach: laparoscopy or robot (assisted) = 0, laparotomy = 1.

Abbreviations: ASA: American Society of Anesthesiologists classification, mILAS: modified Iowa level of assistance scale, NPV: negative predictive value, PPV: positive predictive value, SRT: steep ramp test, WR_{peak}: work rate at peak exercise.

models are nearly equal, leading to a possibly comparable risk prediction for developing complications and the risk of a prolonged in-hospital physical recovery time within the same patient. However, because the individual factors contribute differently to the predicted risk in each model, and in the light of an improved and more patient-oriented information provision about the expected postoperative recovery, using two separate risk predictions might improve the understanding in patients why preoperatively improving physical fitness is important. The impact of complications is not always understood by patients in the preoperative setting, whereas the time needed to regain independent physical functioning postoperatively might appeal more to a patient's imagination [34]. In addition, this risk score for prolonged in-hospital recovery of physical functioning might be used to direct and guide both exercise prehabilitation and postoperative physical therapy management for the patient to accelerate postoperative recovery, and to timely make arrangements regarding postoperative support, either in-hospital or at home after discharge, as well as regarding discharge location.

4.1. Strengths and limitations

Strengths of the current study were a large sample size and a low number of missing values in both cohorts. Furthermore, the model was developed from an academic hospital cohort and externally validated on a cohort with patients from three general hospitals. The academic population differed slightly from the general hospital patients showing statistically significant differences in BMI, ASA classification, preoperative HB, neoadjuvant treatment, and tumour location. Based on the absolute numbers, only the differences in neoadjuvant treatment and tumour location seemed clinically relevant. Despite these differences, the models seemed able to distinguish between patients with and without an increased risk for postoperative complications and a prolonged time to in-hospital physical recovery after elective CRC surgery in both academic and non-academic settings. The study also had several limitations. The study only incorporated preoperative aerobic fitness as modifiable risk factor in predicting postoperative morbidity. Other modifiable risk factors such as nutritional and psychological status were not recorded but might also be of importance in preoperative risk prediction [24,35,36]. Additionally, the observed differences between the derivation and the validation cohort, especially the small percentage of patients who received neoadjuvant treatment and the lower percentage of postoperative complications in the validation cohort, might have influenced the calibration of both models. Future studies should focus on predicting abilities of additional modifiable risk factors and on developing prediction models including multiple modifiable risk factors to further improve patient-tailored preoperative risk prediction, as well as to better target multimodal prehabilitation interventions to the needs of a patient in the attempt to lower the risk of postoperative morbidity.

4.2. Conclusion

This study developed and externally validated two clinical models to preoperatively distinguish between patients with and without an increased risk for postoperative complications, and between patients with and without an increased risk for a prolonged time to in-hospital recovery of physical functioning after elective CRC surgery. Both clinical prediction models can be used for patient-tailored preoperative risk assessment, targeted and cost-effective application of prehabilitation interventions, and improved preoperative patient information regarding the expected course of postoperative in-hospital recovery.

Ethics approval statement and patient consent statement

The Medical Ethical Committee of Maastricht University Medical Centre/Maastricht University approved the study (registration number 15-4-234). Therefore, the study has been performed in accordance with the ethical standards as laid down in the Declaration of Helsinki and its later amendments. Data from all consecutive patients who signed informed consent to use their usual care data for research purposes were prospectively recorded.

Permission to reproduce material from other sources

All authors declare that the submitted work has not been published before (neither in English nor in any other language) and that the work is not under consideration for publication elsewhere.

Financial support

The collaboration project (PROCLINA) was co-funded by an unconditional research grant from MRDM (Medical Research Data

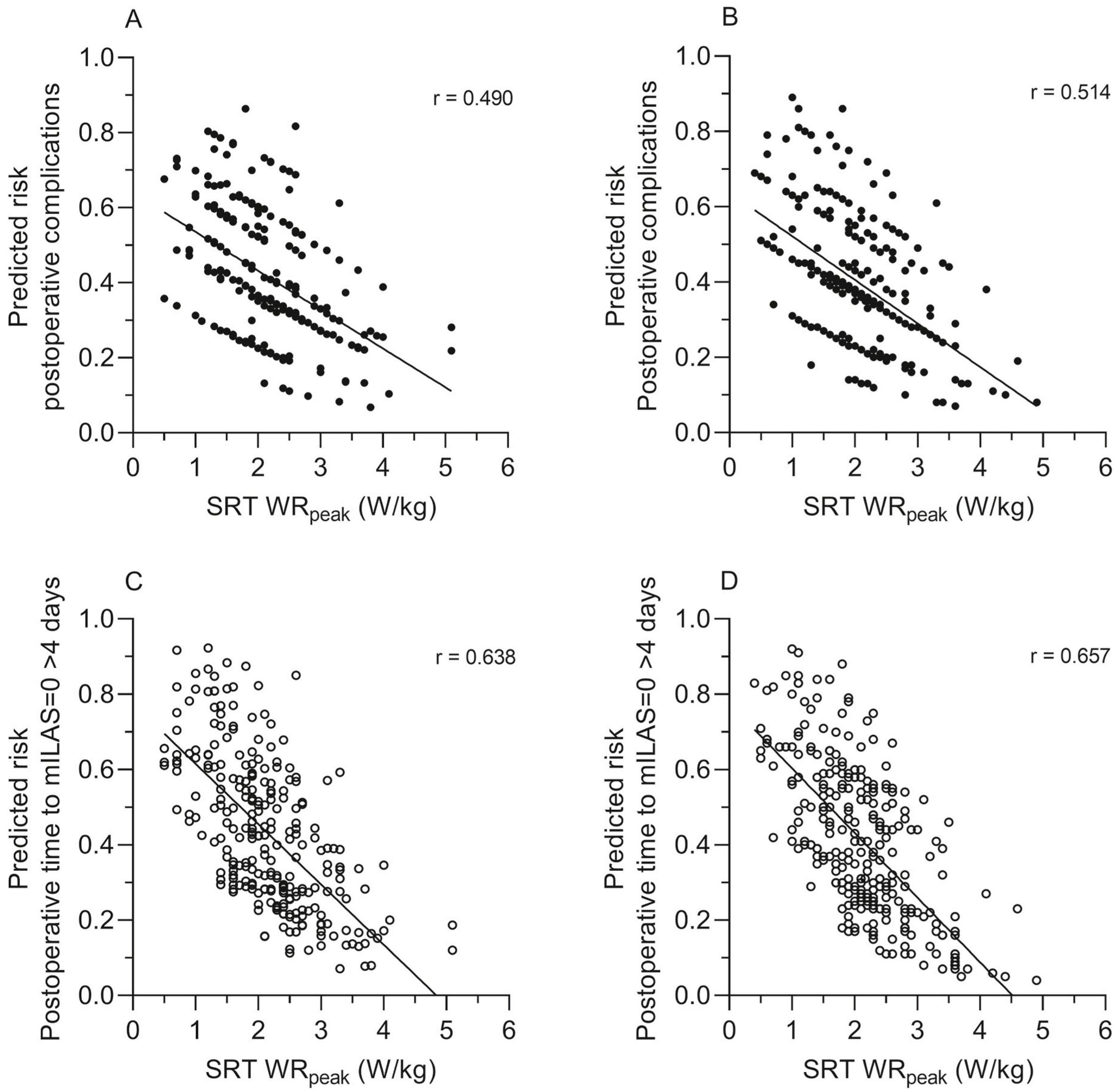

Fig. 4. Preoperative aerobic fitness inversely related to the predicted risk for postoperative complications in the derivation cohort (graph A) and the validation cohort (graph B), and to the predicted risk for a postoperative time to mILAS $= 0 > 4$ days in the derivation cohort (graph C) and the validation cohort (graph D).

Management), as well as by the Ministry of Economic Affairs by means of a PPP Allowance made available by Health~Holland, Top Sector Life Sciences & Health (LSHM17073), to stimulate public-private partnerships. The funding sources had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Funding statement

This collaboration project (PROCLINA) was co-funded by an unconditional research grant from MRDM (Medical Research Data Management), as well as by the Ministry of Economic Affairs by means of a PPP Allowance made available by Health~Holland, Top Sector Life Sciences & Health (LSHM17073). Health~Holland encourages innovative research by financially supporting public-private partnerships in the life sciences and health sector, with the aim of developing sustainable and innovative products and services. The consortium has made agreements about the intellectual property (IP) related to the knowledge and products that will be developed in the project. The funding sources had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of competing interest

The authors declare no conflict of interest.

NvM is professor and executive director of Health~Holland. No staff member of Health~Holland (including the executive director) can ever be involved in the assessment, allocation, and board decisions regarding applications. Health~Holland does not interfere in any way during the implementation of projects. Only after financial and administrative completion of the project, and after delivery of the formal report to Health~Holland, NvM became involved in the writing and editing of this article. Therefore, the authors declare no conflict of (financial) interest.

Acknowledgements

Not applicable.

References

[1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer J Clin 2018;68(6):394–424.
[2] Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer 2019;144(8):1941–53.
[3] Dronkers J, Witteman B, van Meeteren N. Surgery and functional mobility: doing the right thing at the right time. Tech Coloproctol 2016;20(6):339–41.
[4] Hulzebos EH, van Meeteren NL. Making the elderly fit for surgery. Br J Surg 2016;103(2):e12–5.
[5] Etzioni DA, Liu JH, Maggard MA, Ko CY. The aging population and its impact on the surgery workforce. Ann Surg 2003;238(2):170–7.
[6] Rajabiyazdi F, Alam R, Pal A, et al. Understanding the meaning of recovery to patients undergoing abdominal surgery. JAMA Surg 2021;156(8):758–65.
[7] Heldens A, Bongers BC, Lenssen AF, et al. The association between performance parameters of physical fitness and postoperative outcomes in patients undergoing colorectal surgery: an evaluation of care data. Eur J Surg Oncol 2017;43(11):2084–92.
[8] West MA, Asher R, Browning M, et al. Validation of preoperative cardiopulmonary exercise testing-derived variables to predict in-hospital morbidity after major colorectal surgery. Br J Surg 2016;103(6):744–52.
[9] Moran J, Wilson F, Guinan E, et al. Role of cardiopulmonary exercise testing as a risk-assessment method in patients undergoing intra-abdominal surgery: a systematic review. Br J Anaesth 2016;116(2):177–91.
[10] Cuijpers ACM, Heldens A, Bours MJL, et al. Relation between preoperative aerobic fitness estimated by steep ramp test performance and postoperative morbidity in colorectal cancer surgery: prospective observational study. Br J Surg 2022;109(2):155–9.
[11] Barberan-Garcia A, Ubré M, Roca J, et al. Personalised prehabilitation in high-risk patients undergoing elective major abdominal surgery: a randomized blinded controlled trial. Ann Surg 2018;267(1):50–6.
[12] Thomas G, Tahir MR, Bongers BC, et al. Prehabilitation before major intra-abdominal cancer surgery: a systematic review of randomised controlled trials. Eur J Anaesthesiol 2019;36(12):933–45.
[13] Berkel AEM, Bongers BC, Kotte H, et al. Effects of community-based exercise prehabilitation for patients scheduled for colorectal surgery with high risk for postoperative complications: results of a randomized clinical trial. Ann Surg 2022;275(2):e299–306.
[14] Bongers BC. Steep ramp test protocol for preoperative risk assessment and short-term high-intensity interval training to evaluate, improve, and monitor cardiorespiratory fitness in surgical oncology. J Surg Oncol 2023;127(5):891–5.
[15] De Backer IC, Schep G, Hoogeveen A, et al. Exercise testing and training in a cancer rehabilitation program: the advantage of the steep ramp test. Arch Phys Med Rehabil 2007;88(5):610–6.
[16] Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Br Med J 2015;350:g7594.
[17] Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162(1):W1–73.
[18] Gustafsson UO, Scott MJ, Hubner M, et al. Guidelines for perioperative care in elective colorectal surgery: enhanced recovery after surgery (ERAS(®)) society recommendations: 2018. World J Surg 2019;43(3):659–95.
[19] Van Beijsterveld C, Bongers BC, Den Dulk M, et al. Exploring the relation between preoperative physical functioning and the impact of major complications in patients following pancreatic resection. HPB (Oxford) 2020;22(5):716–27.
[20] Van Beijsterveld CA, Bongers BC, Den Dulk M, et al. The association between preoperative physical functioning and short-term postoperative outcomes: a cohort study of patients undergoing elective hepatic resection. HPB (Oxford) 2019;21(10):1362–70.
[21] Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240(2):205–13.
[22] Shields RK, Enloe LJ, Evans RE, et al. Reliability, validity, and responsiveness of functional tests in patients with total joint replacement. Phys Ther 1995;75(3):169–76.
[23] Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49(12):1373–9.
[24] Huisman DE, Reudink M, van Rooijen SJ, et al. LekCheck: a prospective study to identify perioperative modifiable risk factors for anastomotic leakage in colorectal surgery. Ann Surg 2022;275(1):e189–97.
[25] Steyerberg E. Clinical prediction models: a practical approach to development, validation, and updating, vol. 19; 2009.
[26] Snowden CP, Prentis JM, Anderson HL, et al. Submaximal cardiopulmonary exercise testing predicts complications and hospital length of stay in patients undergoing major elective surgery. Ann Surg 2010;251(3):535–41.
[27] West MA, Lythgoe D, Barben CP, et al. Cardiopulmonary exercise variables are associated with postoperative morbidity after major colonic surgery: a prospective blinded observational study. Br J Anaesth 2014;112(4):665–71.
[28] Older P, Hall A, Hader R. Cardiopulmonary exercise testing as a screening test for perioperative management of major surgery in the elderly. Chest 1999;116(2):355–62.
[29] Snowden CP, Prentis J, Jacques B, et al. Cardiorespiratory fitness predicts mortality and hospital length of stay after major elective surgery in older people. Ann Surg 2013;257(6):999–1004.
[30] Lee CHA, Kong JC, Ismail H, et al. Systematic review and meta-analysis of objective assessment of physical fitness in patients undergoing colorectal cancer surgery. Dis Colon Rectum 2018;61(3):400–9.
[31] Nutt CL, Russell JC. Use of the pre-operative shuttle walk test to predict morbidity and mortality after elective major colorectal surgery. Anaesthesia 2012;67(8):839–49.
[32] Weemaes ATR, Beelen M, Bongers BC, et al. Criterion validity and responsiveness of the steep ramp test to evaluate aerobic capacity in survivors of cancer participating in a supervised exercise rehabilitation program. Arch Phys Med Rehabil 2021;102(11):2150–6.
[33] Karlsson E, Egenvall M, Farahnak P, et al. Better preoperative physical performance reduces the odds of complication severity and discharge to care facility after abdominal cancer resection in people over the age of 70 - a prospective cohort study. Eur J Surg Oncol 2018;44(11):1760–7.
[34] Agasi-Idenburg CS, Zuilen MK, Westerman MJ, et al. "I am busy surviving" - views about physical exercise in older adults scheduled for colorectal cancer surgery. J Geriatr Oncol 2020;11(3):444–50.
[35] McDermott FD, Heeney A, Kelly ME, et al. Systematic review of preoperative, intraoperative and postoperative risk factors for colorectal anastomotic leaks. Br J Surg 2015;102(5):462–79.
[36] van Rooijen S, Carli F, Dalton SO, et al. Preoperative modifiable risk factors in colorectal surgery: an observational cohort study identifying the possible value of prehabilitation. Acta Oncol 2017;56(2):329–34.