Acta Oncologica

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ionc20

Accurate population-based model for individual prediction of colon cancer recurrence

E. Osterman, J. Ekström, T. Sjöblom, H. Kørner, T. Å. Myklebust, M. G. Guren & B. Glimelius

To cite this article: E. Osterman, J. Ekström, T. Sjöblom, H. Kørner, T. Å. Myklebust, M. G. Guren & B. Glimelius (2021) Accurate population-based model for individual prediction of colon cancer recurrence, Acta Oncologica, 60:10, 1241-1249, DOI: 10.1080/0284186X.2021.1953138

To link to this article: https://doi.org/10.1080/0284186X.2021.1953138

ACTA ONCOLOGICA
2021, VOL. 60, NO. 10, 1241–1249
https://doi.org/10.1080/0284186X.2021.1953138

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Accurate population-based model for individual prediction of colon cancer recurrence

E. Osterman[a,b], J. Ekström[a], T. Sjöblom[a], H. Kørner[c,d], T. Å. Myklebust[e,f], M. G. Guren[g] and B. Glimelius[a]

[a]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden; [b]Department of Surgery, Region Gävleborg, Gävle, Sweden; [c]Institute of Clinical Science, University of Bergen, Bergen, Norway; [d]Department of Gastrointestinal Surgery, Stavanger University Hospital, Stavanger, Norway; [e]Department of Registration, Cancer Registry of Norway, Oslo, Norway; [f]Department of Research and Innovation, Møre and Romsdal Hospital Trust, Ålesund, Norway; [g]Department of Oncology, Oslo University Hospital, Oslo, Norway

ABSTRACT

Background: Prediction models are useful tools in the clinical management of colon cancer patients, particularly when estimating the recurrence rate and, thus, the need for adjuvant treatment. However, the most used models (MSKCC, ACCENT) are based on several decades-old patient series from clinical trials, likely overestimating the current risk of recurrence, especially in low-risk groups, as outcomes have improved over time. The aim was to develop and validate an updated model for the prediction of recurrence within 5 years after surgery using routinely collected clinicopathologic variables.

Material and methods: A population-based cohort from the Swedish Colorectal Cancer Registry of 16,134 stage I–III colon cancer cases was used. A multivariable model was constructed using Cox proportional hazards regression. Three-quarters of the cases were used for model development and one quarter for internal validation. External validation was performed using 12,769 stage II–III patients from the Norwegian Colorectal Cancer Registry. The model was compared to previous nomograms.

Results: The nomogram consisted of eight variables: sex, sidedness, pT-substages, number of positive and found lymph nodes, emergency surgery, lymphovascular and perineural invasion. The area under the curve (AUC) was 0.78 in the model, 0.76 in internal validation, and 0.70 in external validation. The model calibrated well, especially in low-risk patients, and performed better than existing nomograms in the Swedish registry data. The new nomogram's AUC was equal to that of the MSKCC but the calibration was better.

Conclusion: The nomogram based on recently operated patients from a population registry predicts recurrence risk more accurately than previous nomograms. It performs best in the low-risk groups where the risk-benefit ratio of adjuvant treatment is debatable and the need for an accurate prediction model is the largest.

ARTICLE HISTORY
Received 21 April 2021
Accepted 3 July 2021

KEYWORDS
Colon cancer; nomogram; prognosis; recurrence; adjuvant chemotherapy

Background

Colon cancer is one of the most common cancers in the world and the incidence is expected to increase [1]. Huge efforts have been made to improve the prediction of recurrence risk and few, if any, can exceed the predictive power of the American Joint Committee on Cancer, UICC/TNM-system, and supplementary clinical data [2].

Nomograms are the visual representation of equations; however, the possible inputs and solutions are limited to the variables included in the design. If the risk associated with the variables or the overall outcome changes, the accuracy would inevitably decline.

Currently, there are several nomograms available for prediction of recurrence of colon cancer after radical surgery; the thrice updated Memorial Sloan Kettering Cancer Centre (MSKCC) nomogram [3–5], the ACCENT nomogram [6], a nomogram by Hoshino et al. [7], and a nomogram from Special Commission on Cancer (SCOC) data [8]. The original MSKCC nomogram was based on 1320 patients who underwent surgery between 1990 and 2000 and the ACCENT nomogram is based on 15,936 stage III patients included in adjuvant treatment trials between 1989 and 2002. Previous validations of these nomograms relied on data from the same time period and showed a good correlation with clinical outcomes [9,10]. The nomogram by Hoshino et al. is constructed only for stage II patients and is based on data from patients operated 1997–2006 [7], while the SCOC nomogram is based on a selection of patients operated 2006–2007 at several US hospitals [8].

Since 1989 and 2002, when the two most widely used nomograms included the patients, the quality of care has substantially improved [11–14]. The primary goal of adjuvant chemotherapy is to reduce recurrence rates by eradicating

CONTACT E. Osterman erik.osterman@igp.uu.se Department of Surgery, Region Gävleborg, Lasarettsvägen 1, Gävle 80175, Sweden

Supplemental data for this article can be accessed here.

subclinical tumour deposits that eventually will grow to metastases unless the patient dies from other reasons. The current risks of recurrence and mortality are lower than previously observed [15–17]. The importance of some variables has also likely changed, with improved staging, and other variables have emerged that may be considered [18–20]. The generalisability of data from adjuvant trials can also be questioned, as inclusion is selective and biased towards healthier and younger patients while single academic centres may underestimate the recurrence risk. Using population-based materials should offer better generalisability and utility in the clinical setting.

Aims and hypothesis

Our aim was to develop and externally validate a new nomogram estimating the risk of recurrence after radical surgery for colon cancer, to be used at the postoperative visit for evaluating the need for adjuvant therapy, based on recent population data where the quality of care is known and representative of other developed countries. A second aim was to investigate the accuracy of current nomograms using a Swedish population-based cohort. Our hypothesis was that current nomograms overestimate the risk of recurrence, especially in low-risk groups (stage II–III with no/few risk factors), and that an updated population-based prediction model performs better and thus will be more relevant for patients diagnosed and treated today in developed countries.

Method

The study was approved by the regional ethics committee in Uppsala-Örebro, dnr 2013/093/1(Sweden), and the Regional Ethics Committee South-East 2020/91617 (Norway).

Material

The study cohort included 21,008 cases identified in the nationwide Swedish ColoRectal Cancer Registry (SCRCR) with stage I–III colon adenocarcinoma diagnosed 2007–2014. The last follow-up was on 20 February 2020.

A total of 16,134 cases were eligible for model development after the following cases were omitted from the analyses: multiple surgeries for multiple tumours ($n = 1050$), no surgery or only polypectomy ($n = 1296$), non-radical surgery according to the pathology report ($n = 1958$), patients who died within 30 days of surgery ($n = 317$), who received neo-adjuvant treatment ($n = 233$), and cases with missing or negative time to recurrence ($n = 20$). A cohortogram is presented in Figure 1 which includes both the model, internal and external validation data.

Patients were divided into 5 geographical regions based on the reporting hospital codes [21]. To develop the model 11,943 cases from regions 2–5 were used and 4191 cases from region 1 were used for internal validation. An external validation dataset was obtained from the Norwegian Colorectal Cancer Registry (NCCR) at the Cancer Registry of Norway consisting of 12,769 stage II–III patients with the same selection criteria as the model cohort.

The guidelines for transparent reporting of a multivariable prediction model for individual prognosis or diagnosis and the American Joint Committee on Cancer recommendations were followed [22,23]. The outcome was defined as time to recurrence (TTR) and evaluated as recurrence within 5 years from the date of surgery [24]. Variables from the SCRCR were requested to cover the presently available nomograms and risk factors used in ESMO and NCCN guidelines [3,6–8,25,26]. Variables were: age, sex, patient height, and weight, American Society of Anaesthesiologists (ASA) classification, postoperative complications (yes/no), tumour side (right: caecum to transverse colon, left: splenic flexure to recto-sigmoid junction), emergency surgery (no, perforation, obstruction, other), perioperative perforation (no, yes, peritumoural), T-classification and subclassifications (pT1, pT2, pT3ab; <5 mm of subserosal invasion, pT3cd; ≥ 5 mm of subserosal invasion, pT4a; peritoneal engagement, pT4b; invading other organs), N-classification (pN0-2b), number of found and positive nodes, lymph node ratio (LNR), malignancy grade (high/low), vascular invasion (yes/no), perineural invasion (yes/no), and if referred for or received adjuvant treatment (no, monotherapy, combination therapy), dates of surgery, recurrent disease, and last follow-up or death. Patients who received treatment for recurrent disease, although registered as non-recurrent, were considered recurrent ($n = 6$). The adjuvant

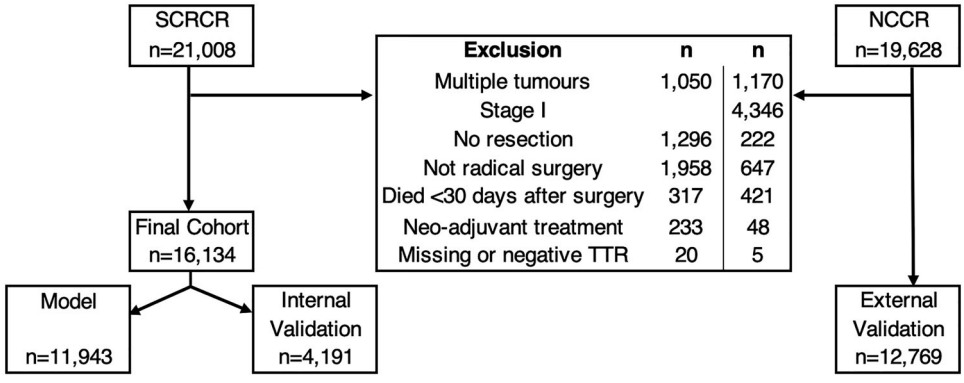

Figure 1. Cohortogram. The Swedish ColoRectal Cancer Registry (SCRCR) data was used to develop the model. Data from the Norwegian Colorectal Cancer Registry (NCCR) was used for external validation.

variable was categorised as 'yes/no' and pT stage 1 and 2 was merged for the multivariable model. Perineural invasion and pT3 and pT4 substage were not available in the Norwegian data, for these patients perineural invasion was set to 'no' and pT3 and pT4 patients were assigned hazards equalling pT3ab patients and pT4a respectively when the external validation was performed.

Statistical analyses

Cox proportional hazards regressions were used to construct the model. Proportional hazards were confirmed by visual inspection of scaled Schoenfeld residuals. To capture non-linearities continuous variables were modelled using restricted cubic splines (RCS) with knots selected by the statistical software for age at 58, 74, and 85, and by a clinical judgement for found nodes at 6, 12, 20, positive nodes at 1, 3, 7, and LNR at 0.01, 0.125, 0.5.

Unadjusted HRs for recurrence were calculated for all variables extracted from the SCRCR, including adjuvant therapy. Variables with clinical and statistical relevance (HR <0.9 or >1.1 or $p < 0.05$ in univariable analyses) were then included in a multivariable model to estimate adjusted HRs. Variables contributing little with regards to HR and significance in the multivariable model were removed to make the nomogram easier to evaluate. Effect plot, calibration curve and area under the curve (AUC) were calculated for the model and validation data separately. Sensitivity was analysed by comparing estimates from models calculated with pT instead of pT-substage, excluding patients who received adjuvant chemotherapy, and excluding patients >80 years old.

Leave-one-out cross-validation was performed with the regions, that is, the model was recalculated five times with all but one region which was then used for validation. The nomogram was drawn from the original multivariable model constructed using regions 2–5 adjusted to no adjuvant treatment. The factor with the highest HR was assigned 100 points, and the other factors assigned points in accordance to the HR in comparison to the highest HR. The risk of recurrence was evaluated at 5 years. Additional axes were drawn for estimated recurrence risk accounting for adjuvant treatment effect with a 25% reduction of recurrences for 5-FU or capecitabine and a further 20% reduction of recurrences with the addition of oxaliplatin for a total reduction of 40% based on data from randomised trials [26–29].

All SCRCR patients were scored retrospectively using the 2008 MSKCC and SCOC nomograms, while only patients treated with adjuvant chemotherapy were scored with the ACCENT nomogram, and stage II patients with the Hoshino nomogram. Since the ACCENT and MSKCC nomogram use preoperative CEA and this was not available in our material we assigned it according to previous research instead of disregarding it as a variable. Thus, 50% of the patients with a recurrence and 35% of non-recurrent patients were randomly assigned an increased CEA value according to previously seen distributions. ASA-classification was substituted for performance status (PS0 = ASA 1–2, PS1 = ASA3, PS2 = ASA $\geq$ 4) in the ACCENT nomogram [30]. The function for recurrence risk from the total score was either estimated from the published nomograms (MSKCC, ACCENT, and Hoshino) or calculated from fictional patients in the online calculator and the corresponding linear predictor from the published model (SCOC). The current nomograms using all SCRCR data and the new nomogram using internal and external validation data were compared using AUC, and the calibration curves (predicted vs observed outcome) drawn using locally estimated scatterplot smoothing to allow for changes in calibration in low and high-risk patients. The area between the calibration curves and optimal prediction was calculated in the range of 10–50% recurrence risk to compare calibration between models.

All Statistical analyses were performed with R 3.6.1 [31] using the packages rms [32] and survival [33].

Results

Table 1 details demographic and tumour characteristics and compares the distribution of recurrences across these characteristics. Male sex, left-sided tumours, emergency surgery, intraoperative perforation, higher pT stage, pN stage, numbers of found and positive nodes, malignancy grade, and vascular and perineural invasion were associated with recurrences. Age below 40 or above 80 years was associated with fewer recurrences. More recurrences in both stage II and III were seen in patients initiating adjuvant treatment. ASA classification and mucinous histology were not associated with recurrences (Supplementary Table 1). Recurrence rates in the cohort increased with the sum of ESMO/NCCN risk factors as presented in Supplementary Table 2. The five-year recurrence rate was 4% in stage I, 11% in stage II, and 27% in stage III.

Model development and validation

Unadjusted HRs and confidence intervals are presented in Supplementary Table 3. The multivariable model was based on previous research, guidelines, known collinearity (e.g., for positive nodes and pN-stage), and the univariable analyses. Several ways to represent node status and risk factors were assessed in adjusted models. Perforation and malignancy grades did not contribute much to the model and were removed in favour of simplicity, HR 0.97 and 0.89 with non-significant p-values. Adjuvant therapy was kept in the model to allow for adjustment. The multivariable model was calculated from 7048 cases after exclusion of patients with missing data (mainly pT3 substage, vascular and perineural invasion). It performed well when tested in model data and is presented in Table 2. The model includes restricted cubic splines to model non-linear effects of both found and positive nodes and the increased risk of missed nodes with low node yields without introducing categorical variables. The non-linear effects are visualised in Supplementary Figure 1 in the form of effect plots. The calibration curves for the model, internal and external validation and other nomograms are presented in Figure 2.

Table 1. Demographics comparison between non-recurrent and recurrent cases in the SCRCR.

Factor	Level	Total	No recurrence	Recurrence	*p*
Sex	Male	7866 (49%)	6477 (82%)	1389 (18%)	<0.0001
	Female	8268 (51%)	7056 (85%)	1212 (15%)	
Side	Right	9270 (57%)	7853 (85%)	1417 (15%)	0.0008
	Left	6854 (42%)	5671 (83%)	1183 (17%)	
	Missing	10 (0%)	9 (90%)	1 (10%)	
Emergency surgery	Elective	14,226 (88%)	12,161 (85%)	2065 (15%)	<0.0001
	Perforation	211 (1%)	162 (77%)	49 (23%)	
	Obstruction	1471 (9%)	1020 (69%)	451 (31%)	
	Other	226 (1%)	190 (84%)	36 (16%)	
pT substage	T1	1072 (7%)	1031 (96%)	41 (4%)	<0.0001
	T2	2399 (15%)	2234 (93%)	165 (7%)	
	T3ab	4396 (27%)	3895 (89%)	501 (11%)	
	T3cd	2615 (16%)	2043 (78%)	572 (22%)	
	T3NA	3240 (20%)	2702 (83%)	538 (17%)	
	T4a	1172 (7%)	769 (66%)	403 (34%)	
	T4b	859 (5%)	579 (67%)	280 (33%)	
	T4NA	328 (2%)	233 (71%)	95 (29%)	
	Missing	53 (0%)	47 (89%)	6 (11%)	
pN substage	N0	10,303 (64%)	9404 (91%)	899 (9%)	<0.0001
	N1a	1860 (12%)	1485 (80%)	375 (20%)	
	N1b	1817 (11%)	1385 (76%)	432 (24%)	
	N2a	1145 (7%)	735 (64%)	410 (36%)	
	N2b	963 (6%)	488 (51%)	475 (49%)	
	Missing	46 (0%)	36 (78%)	10 (22%)	
Positive nodes	Median (IQR)	0.0 (0.0–1.0)	0.0 (0.0–1.0)	2.0 (0.0–5.0)	<0.0001
	Missing	46 (0.3%)	36 (0.3%)	10 (0.4%)	
Found nodes	Median (IQR)	18 (13–25)	18 (13–25)	18 (13–25)	0.17
	Missing	152 (1%)	135 (1%)	17 (1%)	
Lymph node ratio	Median (IQR)	0.0 (0.0–0.1)	0.0 (0.0–0.1)	0.1 (0.0–0.3)	<0.0001
	Missing	45 (0.3%)	34 (0.3%)	11 (0.4%)	
Malignancy grade	Low	12,163 (75%)	10,309 (85%)	1854 (15%)	<0.0001
	High	3174 (20%)	2543 (80%)	631 (20%)	
	Missing	797 (5%)	681 (85%)	116 (15%)	
Vascular invasion	No	11,358 (70%)	9985 (88%)	1373 (12%)	<0.0001
	Yes	3221 (20%)	2234 (69%)	987 (31%)	
	Missing	1555 (10%)	1314 (85%)	241 (15%)	
Perineural invasion	No	11,576 (72%)	10,002 (86%)	1574 (14%)	<0.0001
	Yes	1532 (9%)	1005 (66%)	527 (34%)	
	Missing	3026 (19%)	2526 (83%)	500 (17%)	
Adjuvant therapy	No	11,479 (74%)	10,145 (88%)	1912 (16%)	<0.0001
	Mono	2799 (17%)	2045 (84%)	754 (27%)	
	Combination	1856 (12%)	945 (72%)	513 (28%)	

Note. Fisher's Exact Test was used for categorical variables. Wilcoxon test for continuous variables.

The AUC was 0.78 (95% CI 0.76–0.79) in the model data and 0.76 (95% CI 0.74–0.78) in the internal validation data.

No large deviation from the validation or model data was seen when it was further investigated with leave-one-out cross-validation of the predicted risk based on the regions, Supplementary Figure 2. A nomogram was drawn for risk of recurrence and also adjusting for receipt of adjuvant treatment and is presented in Figure 3. By totalling the points for each characteristic, the nomogram predicts risk of recurrence for the individual patient and the expected risk after adjuvant chemotherapy. Missing pT-substage caused the most exclusions in the multivariable model since data for almost all patients was available concerning pT-stage, but not for the subclassification, sensitivity analysis was performed by substituting pT-substage with pT-stage which did not yield differing estimates for the other variables (Supplementary Table 4). The model was also calculated excluding patients who received chemotherapy, the only marked difference was an increase in the HR estimate for pT4b tumours from 2.1 to 2.8. No difference in model coefficients was seen when the oldest patients were excluded.

Comparison to other nomograms and external validation

The SCRCR nomogram was applied to the Norwegian data (demographics in Supplementary Table 5) with the calibration plot presented in Figure 2. AUC was 0.70 (95% CI 0.69–0.72). Patients from the SCRCR were then scored using the SCOC, MSKCC, ACCENT, and Hoshino nomograms, and the predicted risk of recurrence was compared to the observed outcomes and added to Figure 2. AUC values with 95% CI in parenthesis for the SCOC, MSKCC, ACCENT and Hoshino nomograms were: 0.76 (0.74–0.77), 0.78 (0.76–0.79), 0.72 (0.71–0.73) and 0.69 (0.68–0.72), respectively. The ACCENT and Hoshino nomogram had significantly lower AUC than the other nomograms (Supplementary Figure 3). Compared to model data all but MSKCC had worse AUC. The area between the optimal calibration in the range of 10–50% recurrence risk and calibration curve were 0.07, 0.06, and 0.11 in the model, internal and external validation (closer to 0 is better). The average distance from the optimal prediction for SCOC, MSKCC, ACCENT, and Hoshino nomograms were 0.13, −0.07, −0.10, and −0.19. The SCRCR nomogram

Table 2. The multivariable risk prediction model.

Factor	HR	95% CI	p
Sex			
Female vs male	0.86	0.76–0.97	0.0014
Side			
Left vs right	1.12	0.99–1.27	0.0316
pT			
pT1-2 vs pT3ab	0.62	0.5–0.77	<0.0001
pT3cd vs pT3ab	1.42	1.2–1.68	<0.0001
pT4a vs pT3ab	2.24	1.87–2.68	<0.0001
pT4b vs pT3ab	2.11	1.69–2.63	0.2318
Positive Nodes			
rcs1	1.49	1.41–1.56	<0.0001
rcs2	0.70	0.66–0.74	<0.0001
Found Nodes			
rcs1	0.94	0.91–0.97	<0.0001
rcs2	1.04	1.01–1.07	<0.0001
Emergency Surgery			
Yes vs No	1.51	1.3–1.75	0.0003
Vascular Invasion			
Yes vs No	1.25	1.08–1.44	<0.0001
Perineural Invasion			
Yes vs No	1.36	1.17–1.58	<0.0001
Adjuvant			
Yes vs No	0.95	0.82–1.09	0.303

Variables significant in univariable analysis were initially included. Grade and mucinous histology did not remain significant and were removed from the model. Adjuvant treatment was kept in the model to allow for adjustment in the nomogram. Continuous variables were modelled using restricted cubic splines (rcs) to allow for non-linear effects.
HR: hazard ratio; 95% CI: 95% confidence intervals; p: Wald P test for categorical factors.
For continuous variables modelled using restricted cubic splines (RCS) overall Wald *p*-values are presented instead of the coefficient for each knot. Effect plot for positive and found nodes is presented in Supplementary Figure 1.

gives better predictions than the MSKCC nomogram in most scenarios (see Figure 2, CIs are only overlapping at about 10% recurrence risk).

Discussion

The recurrence risk in patients from the Swedish national population operated between 2007 and 2014 is lower than what has been reported in the original MSKCC study but similar to what was reported from the SCRCR in 2018 [3,16]. Almost 29,000 patients were included in the model and validation making it, to our knowledge, the largest material used so far in the development of a colon cancer nomogram. The data was also more recent than the other nomograms except the 2021 MSKCC nomogram. When comparing the various AUC's, no large differences were seen between the nomograms, but AUC is an insufficient measure if the aim is to support treatment decisions; AUC measures the adequacy of the ranking of patients according to risk, not the accuracy of the prediction which is better gauged from the calibration plot where the model and internal- and external-validation are closer to the optimal line than the other nomograms. Thus, the SCRCR nomogram based on recent population-based data provides improved risk predictions compared to previous nomograms when calibration is taken into account while the stratification of patients seems similar for the SCRCR nomogram, SCOC, and MSKCC nomograms. The MSKCC and Hoshino nomograms provide good predictions in patients with 10–20% recurrence risk but then underestimate the risk. The ACCENT nomogram overestimates the risk in low-risk populations. This could be an effect of the specific populations, specialised centres (MSKCC), adjuvant therapy trials (ACCENT), or that our assumptions regarding CEA and PS were not correct. The Hoshino and MSKCC nomograms include CEA which was not available in our material and was assigned according to the distribution seen in other materials, introducing uncertainty. This may in part explain the underestimation of risk as preoperative CEA contributes a large part of the predictive power in the two nomograms, although it was not predictive in a previous Swedish material [20] and if normalised postoperatively did not confer worse prognosis in an American material [34]. The ACCENT nomogram includes performance status and was here substituted with ASA, but the correlation is not complete, introducing bias in the estimates for ACCENT [30]. However, PS contributes only a small part of the total number of points in the ACCENT nomogram, at most 18 points out of 255 possible which does not explain the difference in calibration. It is reasonable to believe that PS may be relevant in trials where adjuvant chemotherapy is provided to all patients whereas PS is probably not relevant for the risk of having residual subclinical tumour cells after radical surgery. The alternative would have been to exclude these nomograms or to ignore the variable which would likely have disadvantaged these nomograms more [20]. Despite these limitations, we think that our approach is justified for the purpose of the study. The SCOC nomogram is the nomogram most similar to the nomogram we present here, no assumptions were made for the SCOC nomogram which overestimates the risk by 10% in low-risk patients and by 20% in patients with the highest risk.

Our aim was to develop a nomogram that performs well in all risk groups but particularly well in the low-risk groups where the risk-benefit ratio of adjuvant treatment is debatable. The SCRCR nomogram calibrates well in low-risk patients and performs well in patients with up to about 40% risk of recurrence. It performs worse in high-risk patients, where the need for a clinical decision aid is less since adjuvant therapy is routinely indicated in all patients unless contraindications are present, but still at least as good as the other available nomograms. Shared decision-making is still equally important for these patients. The deviation in high-risk patients may be due to uncertainties in compensating for the adjuvant chemotherapy effects in these patients, a dilemma present in all recent patient materials where adjuvant therapy is part of the routine for most patients, or possibly because of regression to the mean. The low- and medium-risk patients show less deviation from the optimal line. Here, the estimates are more precise which may be because the use of adjuvant chemotherapy in stage II patients have been rather restricted in Sweden. Non-registered recurrences may explain a small part of the deviation in high-risk patients; a recent investigation of the SCRCR estimated that 1–2% of recurrences at 5 years are missing in the registry [35].

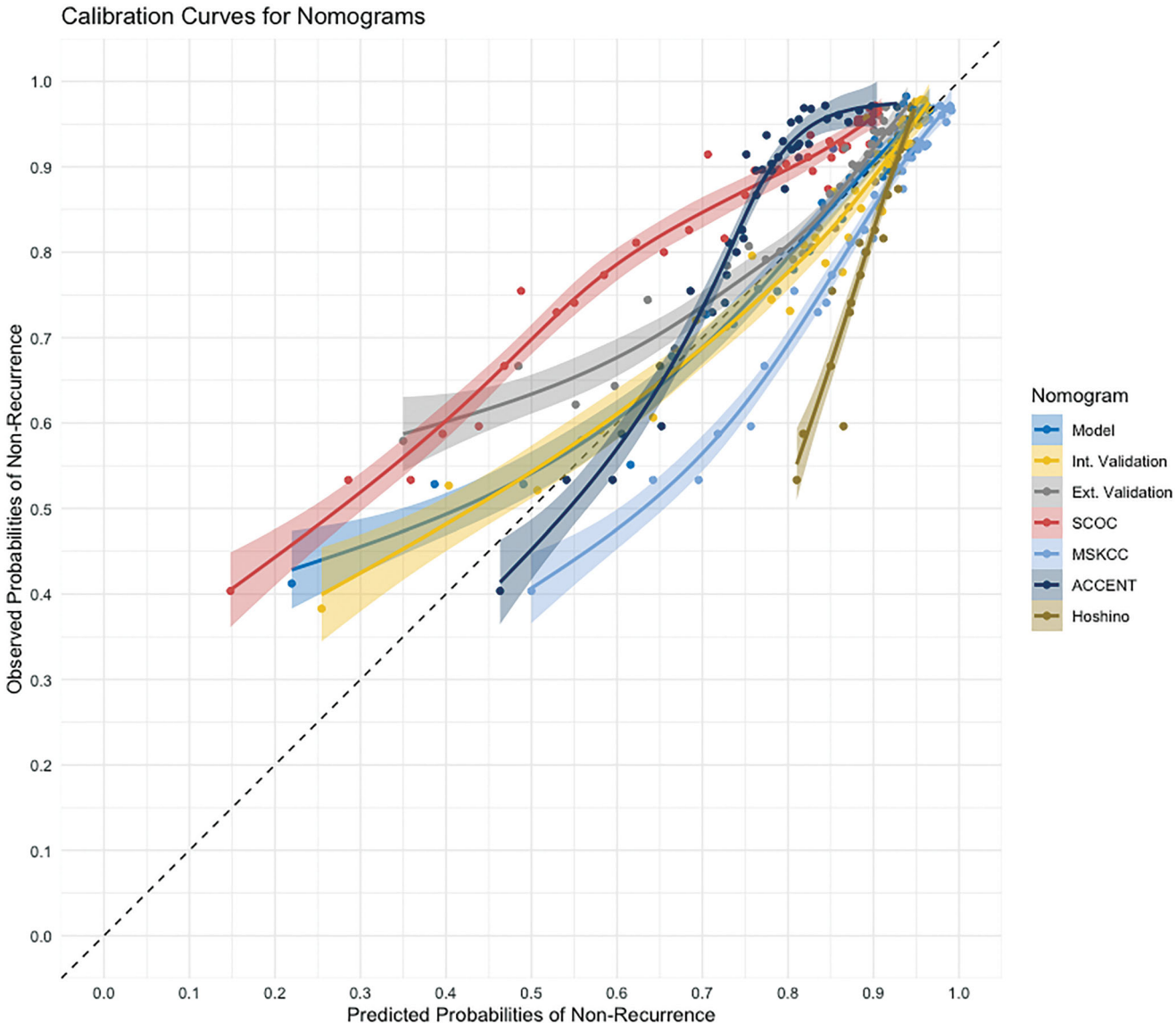

Figure 2. Model calibration curve. Predicted probabilities of non-recurrences vs observed probabilities of non-recurrence in model data, validation data (Int. Validation), external validation in Norwegian data (Ext. Validation) and calibration of other nomograms in SCRCR data (SCOC, MSKCC, ACCENT, Hoshino). Lines were fitted with locally estimated scatterplot smoothing ('loess'). Shaded area denotes 95% confidence intervals.

In the external validation sample, the AUC was 0.70 and similar to how other nomograms performed in the Swedish data. However, the calibration was better in low-intermediate risk patients than the other nomograms but got progressively worse with better-observed prognosis in high-risk patients. This may be an effect of good survival outcomes in Norway [36], regression to the mean or recurrences not reported to the registry (94% conformity between patient records and registry [37], lower than in Sweden [35]). A weakness is that we did not externally validate the predictions in stage I patients, however, it is not difficult to predict the outcome since recurrence rates are very low for most if not all stage I patients.

There are few available registries to validate population-based predictions as most registries do not record recurrences [38]. Thus, SCRCR nomogram may not be applicable outside the Nordic countries and Western Europe where the populations are predominantly Caucasian. The patient material also impacts the comparisons with the Hoshino nomogram which is based on Asian patients. Effects from how the healthcare system is organised and financed might impact the validity in other countries. Mucinous histology and high malignancy grade were not strongly associated with recurrences and excluded from the nomogram even though they have previously been described as risk factors, however not seen in the SCRCR [16]. There were few cases with perioperative perforation which may explain its non-significance. Knowledge about mutations, especially mismatch repair deficiency which are associated with a good prognosis at least in stage II may further improve the accuracy of nomograms [39]. This information was not available in the current material and precluded us from evaluating the predictive power of the new MSKCC nomogram [5]. Other markers, for example, circulating tumour DNA and tumour infiltrating lymphocytes,

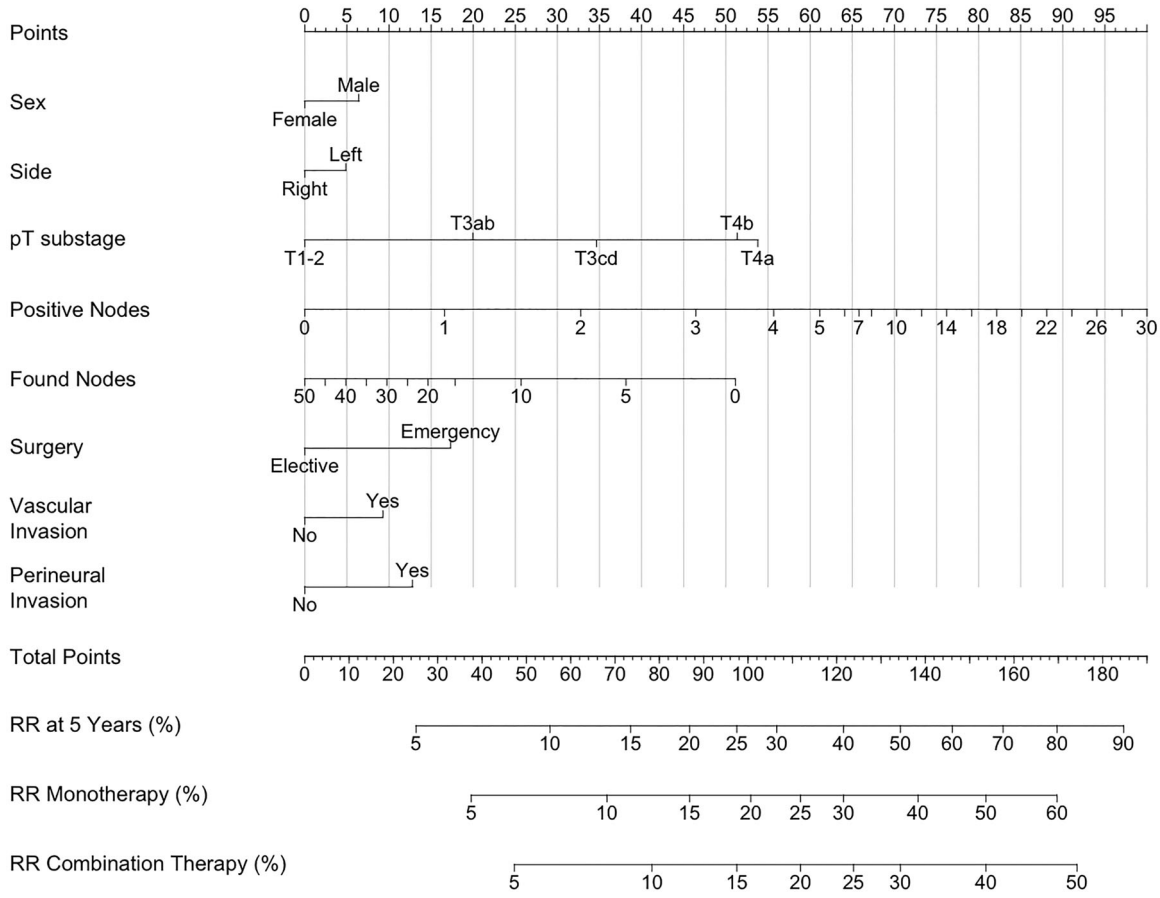

Figure 3. Nomogram. Start from second axis (sex), sum corresponding points read from top axis, continue with each axis until you reach the axis for total points. The total points correspond to the predicted recurrence risk at 5 years. Subsequent axes represent predicted recurrence risk if adjuvant mono- or combination chemotherapy is given. RR: recurrence risk. For example, a male with a right-sided, pT3cd with 2/20 positive nodes and vascular invasion gets 98 points ($7+0+34+33+15+0+9+0$) corresponding to a recurrence risk of about 25%. The next two axes describe the risk of recurrence if adjuvant therapy is given, with monotherapy it is 19% and with combination therapy the recurrence risk is 15.

of increased recurrence risk will likely be of value in the future, though these are seldom available in population-based cohorts.

When the aim is to provide an estimate, stratified for stage and risk factors, that a radically operated patient has remaining tumour cells capable of growing to a recurrence, TTR is the appropriate endpoint. Other outcomes, like disease-free survival, may be more appropriate in clinical trials, as recently discussed by Cohen et al. [40], where gains and losses have to be balanced, but not when deciding the need for adjuvant therapy for the individual patient. Every estimate of recurrence risk according to today's standard of care has to be made in an environment where adjuvant chemotherapy, known to reduce recurrence risks, is provided selectively. This use of adjuvant chemotherapy complicates the calculations and interpretations of data, but much less so in early tumours, in particular stage II, where this is not routine unless risk factors are present. In more advanced tumours (mainly stage III) where adjuvant therapy is provided unless the patient is very old, has co-morbidities, or postoperative complications delaying treatment beyond the recommended time frame of 8–12 weeks, it is more complicated. Age, co-morbidities, and probably also postoperative complications do not have any major influence on whether the tumour has spread outside the tissues normally removed during bowel cancer surgery. Therefore, estimates of the risks in patients not treated with adjuvant treatment likely do not bias TTR.

Conclusions

Since the late 1990s and early 2000s, the standard of care for the cure of colon cancer has improved and the risk of recurrence is lower today. Accordingly, tools for estimation of the recurrence need to be based on contemporary data. The purpose of this study was to develop an up-to-date nomogram supporting the clinician at the postoperative visit in the decision whether to treat with adjuvant chemotherapy or not, instead of solely predicting the prognosis. The SCRCR nomogram developed here is based on resections performed at approximately 50 different hospitals, but with one national care guideline, in a population of about 10 million inhabitants, making it truly population-based. It performs well in low-risk groups where the clinical need for decision aids is the greatest and it validates internally and externally.

Authors' contributions

EO, JE, TS, BG contributed to the conception and design of the work. EO, HK, TÅM, MGG, BG acquired data. EO, JE, TÅG analysed the material. EO, MGG, BG drafted the work. All authors have read and contributed revisions to the work and have approved the submitted version.

Acknowledgments

All rapporteurs and centres participating in the SCRCR and Norwegian Colorectal Cancer Registry. The regional cancer centre, Healthcare region Mid Sweden and Hans Garmo for statistical advice.

Disclosure statement

The authors report no conflict of interest. This work has not been presented elsewhere. The study has used data from the Cancer Registry of Norway. The interpretation and reporting of these data are the sole responsibility of the authors, and no endorsement by the Swedish Colorectal Cancer Registry or the Cancer Registry of Norway is intended nor should be inferred.

Funding

This work was supported by the Swedish Cancer Society [19 0382], Gävle Cancer Society, and by the Norwegian Cancer Society [190188-2017].

ORCID

E. Osterman http://orcid.org/0000-0003-1621-7872
T. Sjöblom http://orcid.org/0000-0001-6668-4140
H. Kørner http://orcid.org/0000-0003-0498-2331
T. Å. Myklebust http://orcid.org/0000-0003-4645-1635
M. G. Guren http://orcid.org/0000-0003-0593-0218
B. Glimelius http://orcid.org/0000-0002-5440-791X

References

[1] Arnold M, Sierra MS, Laversanne M, et al. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683–691.
[2] Puccini A, Berger MD, Zhang W, et al. What we know about stage II and III colon cancer: it's still not enough. Targ Oncol. 2017;12(3):265–275.
[3] Weiser MR, Landmann RG, Kattan MW, et al. Individualized prediction of colon cancer recurrence using a nomogram. J Clin Oncol. 2008;26(3):380–385.
[4] Konishi T, Shimada Y, Hsu M, et al. Contemporary validation of a nomogram predicting colon cancer recurrence, revealing all-stage improved outcomes. JNCI Cancer Spectr. 2019;3(2):pkz015.
[5] Weiser MR, Hsu M, Bauer PS, et al. Clinical calculator based on molecular and clinicopathologic characteristics predicts recurrence following resection of stage I-III colon cancer. J Clin Oncol. 2021;39(8):911–919.
[6] Renfro LA, Grothey A, Xue Y, et al. ACCENT-based web calculators to predict recurrence and overall survival in stage III colon cancer. J Natl Cancer Inst. 2014;106(12):dju333.
[7] Hoshino N, Hasegawa S, Hida K, et al. Nomogram for predicting recurrence in stage II colorectal cancer. Acta Oncol. 2016;55(12):1414–1417.
[8] Zafar SN, Hu C-Y, Snyder RA, et al. Predicting risk of recurrence after colorectal cancer surgery in the United States: an analysis of a special commission on cancer national study. Ann Surg Oncol. 2020;27(8):2740–2749.
[9] Kazem MA, Khan AU, Selvasekar CR. Validation of nomogram for disease free survival for colon cancer in UK population: a prospective cohort study. Int J Surg. 2016;27:58–65.
[10] Collins IM, Kelleher F, Stuart C, et al. Clinical decision aids in colon cancer: a comparison of two predictive nomograms. Clin Colorectal Cancer. 2012;11(2):138–142.
[11] Bokey L, Chapuis PH, Chan C, et al. Long-term results following an anatomically based surgical technique for resection of colon cancer: a comparison with results from complete mesocolic excision. Colorectal Dis. 2016;18(7):676–683.
[12] Hohenberger W, Weber K, Matzel K, et al. Standardized surgery for colonic cancer: complete mesocolic excision and central ligation-technical notes and outcome. Colorectal Dis. 2009;11(4):354–364.
[13] Iversen LH, Green A, Ingeholm P, et al. Improved survival of colorectal cancer in Denmark during 2001–2012 – The efforts of several national initiatives. Acta Oncologica. 2016;55(2):10–23.
[14] Walters S, Benitez-Majano S, Muller P, et al. Is England closing the international gap in cancer survival? Br J Cancer. 2015;113(5):848–860.
[15] White A, Joseph D, Rim SH, et al. Colon cancer survival in the United States by race and stage (2001-2009): findings from the CONCORD-2 study. Cancer. 2017;123:5014–5036.
[16] Osterman E, Glimelius B. Recurrence risk after up-to-date colon cancer staging, surgery, and pathology: analysis of the entire Swedish population. Dis Colon Rectum. 2018;61(9):1016–1025.
[17] Böckelman C, Engelmann BE, Kaprio T, et al. Risk of recurrence in patients with colon cancer stage II and III: a systematic review and meta-analysis of recent literature. Acta Oncol. 2015;54(1):5–16.
[18] Shi Q, Andre T, Grothey A, et al. Comparison of outcomes after fluorouracil-based adjuvant therapy for stages II and III colon cancer between 1978 to 1995 and 1996 to 2007: evidence of stage migration from the ACCENT database. J Clin Oncol. 2013;31(29):3656–3663.
[19] Salem ME, Yin J, Goldberg RM, et al. Evaluation of the change of outcomes over a 10-year period in patients with stage III colon cancer: pooled analysis of 6501 patients treated with fluorouracil, leucovorin, and oxaliplatin in the ACCENT database. Ann Oncol. 2020;31(4):480–486.
[20] Osterman E, Mezheyeuski A, Sjöblom T, et al. Beyond the NCCN risk factors in colon cancer: an evaluation in a Swedish population-based cohort. Ann Surg Oncol. 2020;27(4):1036–1045.
[21] Erik Bulow. Decoder: decode coded variables to plain text and the other way around. R package version 1.2.2. [cited 2020 May 14] 2020. Available from: https://CRAN.R-project.org/package=decoder
[22] Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIP): the TRIPOD statement. Ann Intern Med. 2015;162(1):55–63.
[23] Kattan MW, Hess KR, Amin MB, et al. American Joint Committee on Cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine: AJCC criteria for risk models in precision medicine. CA Cancer J Clin. 2016;66(5):370–374.
[24] Punt CJA, Buyse M, Köhne C-H, et al. Endpoints in adjuvant treatment trials: a systematic review of the literature in colon cancer and proposed definitions for future trials. J Natl Cancer Inst. 2007;99(13):998–1003.
[25] Benson AB, Venook AP, Al-Hawary MM, et al. NCCN guidelines insights: colon cancer, version 2.2018. J Natl Compr Canc Netw. 2018;16(4):359–369.
[26] Labianca R, Nordlinger B, Beretta GD, et al. Early colon cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2013;24:vi64–vi72.
[27] Yothers G, O'Connell MJ, Allegra CJ, et al. Oxaliplatin as adjuvant therapy for colon cancer: updated results of NSABP C-07 trial, including survival and subset analyses. J Clin Oncol. 2011;29(28):3768–3774.

[28] Schmoll H-J, Twelves C, Sun W, et al. Effect of adjuvant capecitabine or fluorouracil, with or without oxaliplatin, on survival outcomes in stage III colon cancer and the effect of oxaliplatin on post-relapse survival: a pooled analysis of individual patient data from four randomised controlled trials. Lancet Oncol. 2014;15(13):1481–1492.

[29] Efficacy of adjuvant fluorouracil and folinic acid in colon cancer: International Multicentre Pooled Analysis of Colon Cancer Trials (IMPACT) investigators. The Lancet. 1995;345:939–944.

[30] Young J, Badgery-Parker T, Dobbins T, et al. Comparison of ECOG/WHO performance status and ASA score as a measure of functional status. J Pain Symptom Manage. 2015;49(2):258–264.

[31] R Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2013.

[32] Harrell FE Jr. rms: Regression Modeling Strategies. R package version 5.1-4. [cited 2020 May 14] 2019. Available from: https://CRAN.R-project.org/package=rms

[33] Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. New York (NY): Springer; 2000.

[34] Konishi T, Shimada Y, Hsu M, et al. Association of preoperative and postoperative serum carcinoembryonic antigen and colon cancer outcome. JAMA Oncol. 2018;4(3):309–315.

[35] Osterman E, Hammarström K, Imam I, et al. Completeness and accuracy of the registration of recurrences in the Swedish Colorectal Cancer Registry (SCRCR) and an update of recurrence risk in colon cancer. Acta Oncologica. 2021;60(7):842–849.

[36] Araghi M, Arnold M, Rutherford MJ, et al. Colon and rectal cancer survival in seven high-income countries 2010-2014: variation by age and stage at diagnosis (the ICBP SURVMARK-2 project). Gut. 2021;70(1):114–126.

[37] Dørum LM, Larsen IK, Knudsen KO, et al. Evaluering av datakvaliteten i Nasjonalt kvalitetsregister for tykk- og endetarmskreft. Oslo: Kreftregistret. 2017. p. 26–28.

[38] MacCallum C, Skandarajah A, Gibbs P, et al. The value of clinical colorectal cancer registries in colorectal cancer research: a systematic review. JAMA Surg. 2018;153(9):841–849.

[39] Ribic CM, Sargent DJ, Moore MJ, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. N Engl J Med. 2003;349(3):247–257.

[40] Cohen R, Vernerey D, Bellera C, et al. Guidelines for time-to-event end-point definitions in adjuvant randomised trials for patients with localised colon cancer: results of the DATECAN initiative. Eur J Cancer. 2020;130:63–71.