Development and validation of a risk prediction model for post-polypectomy colorectal cancer in the USA: a prospective cohort study

Markus Dines Knudsen,[a,b,c] Kai Wang,[a] Liang Wang,[a,d] Georgios Polychronidis,[a,e] Paula Berstad,[b] Kana Wu,[f] Xiaosheng He,[g,h,i] Dong Hang,[f,j] Zhe Fang,[a] Shuji Ogino,[a,k,l] Andrew T. Chan,[g,h,m,n] Edward Giovannucci,[a,f] Molin Wang,[a,m] and Mingyang Song[a,f,g,h,*]

[a]Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA, USA
[b]Section of Bowel Cancer Screening, Cancer Registry of Norway, Ullernchausseen 64, Oslo, Norway
[c]Division of Surgery, Inflammatory Diseases and Transplantation, Department of Transplantation Medicine, Oslo University Hospital, Sognsvannsveien 20, Oslo, Norway
[d]Centre of Gastrointestinal Surgery, The First Affiliated Hospital, Sun Yat-sen University, 58 Zhongshaner Rd, Yuexiu District, Guangzhou, Guangdong Province, China
[e]Department of General Visceral and Transplantation Surgery, Heidelberg University Hospital, Im Neuenheimer Feld 420, Heidelberg, Germany
[f]Department of Nutrition, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA, USA
[g]Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, 55 Fruit St, Boston, MA, USA
[h]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, 55 Fruit St, Boston, MA, USA
[i]Department of Colorectal Surgery, The Six Affiliated Hospital, Sun Yat-sen University, 135, Xingang Xi Road, Guangzhou, China
[j]Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, 101 Longmian Avenue, Nanjing, China
[k]Department of Pathology, Program in Molecular Pathological Epidemiology, Brigham and Women's Hospital and Harvard Medical School, 75 Francis Street, Boston, MA, USA
[l]Broad Institute of MIT and Harvard, Merkin Building, 415 Main St, Cambridge, MA, USA
[m]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 181 Longwood Avenue, Boston, MA, USA
[n]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA, USA

eClinicalMedicine 2023;62: 102139
Published Online xxx
https://doi.org/10.1016/j.eclinm.2023.102139

Summary

Background Effective risk stratification tools for post-polypectomy colorectal cancer (PPCRC) are lacking. We aimed to develop an effective risk stratification tool for the prediction of PPCRC in three large population-based cohorts and to validate the tool in a clinical cohort.

Methods Leveraging the integrated endoscopic, histopathologic and epidemiologic data in three U.S population-based cohorts of health professional (the Nurses' Health Study (NHS) I, II and Health Professionals Follow-up Study (HPFS)), we developed a risk score to predict incident PPCRC among 26,741 patients with a polypectomy between 1986 and 2017. We validated the PPCRC score in the Mass General Brigham (MGB) Colonoscopy Cohort (Boston, Massachusetts, U.S) of 76,603 patients with a polypectomy between 2007 and 2018. In all four cohorts, we collected detailed data on patients' demographics, endoscopic history, polyp features, and lifestyle factors at polypectomy. The outcome, incidence of PPCRC, was assessed by biennial follow-up questionnaires in the NHS/HPFS cohorts, and through linkage to the Massachusetts Cancer Registry in the MGB cohort. In all four cohorts, individuals who were diagnosed with CRC or died before baseline or within six months after baseline were excluded. We used Cox regression to calculate the hazard ratio (HR), 95% confidence interval (CI) and assessed the discrimination using C-statistics and reclassification using the Net Reclassification Improvement (NRI).

Findings During a median follow-up of 12.8 years (interquartile range (IQR): 9.3, 16.7) and 5.1 years (IQR: 2.7, 7.8) in the NHS/HPFS and MGB cohorts, we documented 220 and 241 PPCRC cases, respectively. We identified a PPCRC risk score based on 11 predictors. In the validation cohort, the PPCRC risk score showed a strong association with PPCRC risk (HR for high vs. low, 3.55, 95% CI, 2.59–4.88) and demonstrated a C-statistic (95% CI) of 0.75

*Corresponding author. Department of Epidemiology, Harvard T. H. Chan School of Public Health, 677 Huntington Ave, Boston, MA, 02115, USA.
E-mail address: mis911@mail.harvard.edu (M. Song).

(0.70–0.79), and was discriminatory even within the low- and high-risk polyp groups (C-statistic, 0.73 and 0.71, respectively) defined by the current colonoscopy surveillance recommendations, leading to a NRI of 45% (95% CI, 36–54%) for patients with PPCRC.

Interpretation We developed and validated a risk stratification model for PPCRC that may be useful to guide tailored colonoscopy surveillance. Further work is needed to determine the optimal surveillance interval and test the added value of other predictors of PPCRC beyond those included in the current study, along with implementation studies.

Funding US National Institutes of Health, the American Cancer Society, the South-Eastern Norway Regional Health Authority, the Deutsche Forschungsgemeinschaft.

Keywords: Prediction model; Post-polypectomy; Colonoscopy; Colorectal cancer; Colonoscopy surveillance

Research in context

Evidence before this study

The current recommendations for coloscopy surveillance are based only on the polyp findings, ignoring other well established risk factors for colorectal cancer such as demographic and lifestyle factors. Effective risk stratification tools for post-polypectomy colorectal cancer are lacking to tailor colonoscopy surveillance. We used the 2020 US Multi-Society Task Force (USMSTF) recommendations to investigate the current evidence and recommendations for surveillance colonoscopy. We further used the World Cancer Research Fund report for the most updated evidence on lifestyle risk factors and colorectal cancer. We also searched PubMed for peer-reviewed original articles in English published between January 1, 1976, and November 31, 2022, with keywords "prediction model", "post-polypectomy", "colonoscopy" "colorectal cancer", "colonoscopy surveillance". All of which highlighted a dearth of evidence on this specific topic.

Added value of this study

We created and validated a risk score to predict post-polypectomy colorectal cancer (PPCRC) that performed better than the current clinical guidelines for colonoscopy surveillance in a clinical cohort. The results of this study provide novel insights into PPCRC and a proof of concept for risk-based colonoscopy surveillance based on clinical, demographic, and lifestyle risk factors.

Implications of all the available evidence

Our results indicate the importance of patients' demographic and lifestyle characteristics for risk assessment of post-polypectomy colorectal cancer and should be considered in recommendation making for colonoscopy surveillance. Further work is needed. Further modelling may incorporate information on molecular features of polyps and adenoma/polyp detection rates, as these factors have been shown to be important for predicting risk of PPCRC.

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in both men and women in the United States.[1] Endoscopic screening has been shown to reduce CRC incidence and mortality.[2–6] Patients who received polypectomy at a screening colonoscopy are advised to undergo colonoscopy surveillance at certain intervals, based on the most severe polyp findings, to prevent polyp recurrence and post-polypectomy CRC (PPCRC).[7,8]

Colonoscopy surveillance accounts for 25% of all colonoscopy examinations and consumes more than $4 billion annually in US healthcare expenditure.[9,10] This figure is expected to grow due to the increasing uptake of screening,[11] particularly in younger adults as the recommended starting age of screening has been lowered from 50 to 45 years due to the increasing incidence of CRC in adults younger than 50 years.[12,13] However, there are limited data that colonoscopy surveillance reduces CRC incidence or mortality.[7,14–16] As a result, clinical guidelines for colonoscopy surveillance vary widely and lack sufficient support.[7,17–20] Furthermore, adherence to the surveillance recommendations has been shown to be poor, with the reported prevalence of less than 15%, and there is overuse of surveillance colonoscopy in low-risk patients and underuse in high-risk patients.[21–24] Given the invasiveness, high cost, and possible serious complications of colonoscopy,[25] it is vital to minimise unnecessary surveillance for patients who are at low risk of developing PPCRC and ensure adequate surveillance for patients at high risk. Therefore, effective risk stratification tools are urgently needed to tailor colonoscopy surveillance to prevent PPCRC. Yet, prediction models for risk stratification of colonoscopy surveillance are scarce. A recent study developed a prediction model for metachronous advanced neoplasia and showed that the model including polyp features, major CRC risk factors, and colonoscopy quality measure improved post-polypectomy risk stratification compared to the 2020 US Multi-Society Task Force (USMSTF) recommendations.[26]

Therefore, we aimed to develop an effective risk stratification tool for prediction of PPCRC in three large population-based cohorts, including the Nurses' Health Study (NHS) I/II and Health Professional Follow-up Study (HPFS), and then validate and test the tool in a clinical cohort, the Mass General Brigham (MGB) Colonoscopy Cohort, using the risk categorisation in the USMSTF recommendations as the reference.[7] For the predictors, we considered a variety of demographic, clinical (including index colonoscopy- and polyp-related features), and epidemiologic factors (including major lifestyle factors) that have been associated with PPCRC risk.[27–30]

Methods

Study populations

Details of the study cohorts are described in the Supplementary Materials: study populations (pp. 2). The protocol is also provided in the Supplementary Materials.

NHS I/II and HPFS

The NHS I included 121,700 registered US female nurses aged 30 to 55 at enrolment in 1976. The NHS II included 116,429 registered US female nurses aged 25–42 years at enrolment in 1989. The HPFS enrolled 51,529 male health professionals, aged 40–75 at study entry in 1986. The average follow-up rate has been greater than 90% in all three cohorts. More details have been described previously.[31–33] Briefly, participants were mailed a biennial validated questionnaire that inquired detailed medical and lifestyle information.[34,35] If participants reported a diagnosis of polyp, we asked for permission to acquire their endoscopic and pathologic records. A previous validation study in a random sample of 114 women indicated an accuracy of 97% for self-reported negative endoscopies.[36]

A total of 26,741 participants who had a first-time polypectomy between June 1986 and June 2014 in NHS I, between June 1991 and June 2015 in the NHS II, and between June 1986 and June 2010 in the HPFS were included in the current study. Individuals diagnosed with CRC or died before baseline or within six months after baseline were excluded.

MGB colonoscopy cohort

The MGB Cohort is an electronic health record (EHR)-based clinical cohort that included 213,924 patients who had undergone a colonoscopy in 2007–2018 in an integrated healthcare system that included Brigham and Women's Hospital, Brigham and Women's Faulkner Hospital, and Massachusetts General Hospital, known as the Mass General Brigham in Boston, Massachusetts, USA. Details of the cohort have been described previously.[37] Briefly, we extracted detailed endoscopic, pathologic and family history of CRC data from the Provation® endoscopy reporting system and the Research Patient Data Registry (RPDR).[38] We used the unique medical record number and date of colonoscopy to match the endoscopy and pathology data. Demographic and lifestyle data were obtained from the RPDR.

In the present study, a total of 76,603 patients who had their first-time polypectomy were included in the study.

Predictors

We included three groups of predictors, all measured at the time of polypectomy.

(1) Clinical predictors: index polyp histology (serrated polyps, tubular adenomas, tubulovillous adenomas, villous adenomas, carcinoma in situ/high grade dysplasia), having ≥3 adenomas (yes/no), having ≥3 serrated polyps (yes/no), large serrated polyps (≥10 mm) (yes/no), large adenoma (≥10 mm) (yes/no), anatomic placement of adenomas (rectum, distal colon, proximal colon), anatomic placement of serrated polyps (rectum, distal colon, proximal colon), having ≥1 prior endoscopy (yes/no), family history of CRC in any first-degree relative (yes/no), and reason for index endoscopy (screening/symptoms). In addition, for the MGB cohort only, we included quality measures of index endoscopy, including quality of bowel preparation (excellent, good, fair, poor), referring to how clean/empty the bowel was at endoscopy, and cecal intubation (no/yes) referring to if the last part of the colon was reached.

(2) Demographics predictors: age at index endoscopy (40–44, 45–49, 50–54, 55–59, 60–64, 65–69, ≥70), sex (female/male), and race/ethnicity (white/non-white).

(3) Lifestyle predictors: alcohol >1 drink/day for women, >2 drinks/day for men (yes/no), Body mass index (BMI) < 18.5, 18.5–24.9, 25.0–29.9, ≥30 kg/m^2), smoking (never, past, current), regular use of aspirin/Non-steroidal anti-inflammatory drugs (NSAID) ≥ 2 tablets per week (yes/no), and physical activity ≥7.5 metabolic equivalent task score (METS)-hours per week (yes/no). For the MGB cohort, we included BMI only, because other lifestyle variables either were lacking or had very high missingness.

Outcome assessment

In the NHS/HPFS cohorts, once participants reported a diagnosis of CRC on biennial questionnaires, we asked for their permission to collect and review their medical records. Study investigators blinded to the exposure status reviewed the records to confirm the diagnosis and extract relevant clinical information. Death of a cohort member was notified by the next-of-kin or the post office when questionnaires or newsletters were returned or

identified through search of the vital records of states and of the National Death Index. This method has been shown highly sensitive for identifying deaths in our cohorts, with a sensitivity of ~98% and a specificity of 100%.[39]

In the MGB cohort, incidence of CRC after polypectomy was identified through linkage to the Massachusetts Cancer Registry (MCR).[37] The MCR provides state-wide information on cancer diagnosis, including date and histopathology of CRC. It is by law demanded by hospitals and health personal in Massachusetts to report cancers to the MCR[40] and thus completeness of cancer ascertainment is expected to be high. Death was identified from the hospital records and through linkage to the Social Security Death Index.

Risk classification according to the 2020 USMSTF surveillance recommendations

We classified index polyps according to the 2020 USMSTF surveillance recommendations. High-risk polyps were those with a recommended surveillance interval of three years (including 5–10 tubular adenomas <10 mm, adenoma ≥10 mm, adenoma with tubulovillous or villous histology, adenoma with high-grade dysplasia, >10 adenomas, 5–10 sessile serrated polyps [SSPs] <10 mm, SSPs ≥10 mm, SSPs with dysplasia, hyperplastic polyps [HPs] ≥10 mm and/or, traditional serrated adenomas [TSAs] as their most severe findings), and low-risk polyps were all other polyps (including patients with 1–2 tubular adenomas <10 mm, ≤20 HPs in the rectum or sigmoid colon <10 mm, ≤20 HPs proximal to sigmoid colon <10 mm, 1–2 SSPs <10 mm, 3–4 tubular adenomas <10 mm, or 3–4 SSPs <10 mm as their most severe findings).[7]

Ethics

The NHS/HPFS cohorts were approved by the institutional review board at the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required (Institutional Review Board Protocol 10,372). Participants in the NHS/HPFS gave their informed consent by completing and returning of the questionnaires. The MGB cohort was approved by the MGB Human Research Committee (Institutional Review Board Protocol 2018P002052). Individual informed consent was waived for the MGB as the study was considered secondary analysis of existing data.

Statistical analyses

Participants were followed from six months after index polypectomy until the date of CRC diagnosis,[41] death, or the end of follow-up (June 2016 for NHS I/HPFS, June 2017 for NHS II, and October, 2018 for the MGB cohort), whichever came first. We adopted a six-month window to avoid counting the diagnostic endoscopy for CRC as the index endoscopy. Cox proportional hazards regression model was used for prediction modelling.

We first included all predictors in the model and calculated the multivariable-adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) for each of the predictors in the NHS/HPFS cohorts. We calculated the Uno's C-statistics[42] for the models including each of the predictor categories separately (clinical, demographic, and lifestyle) and then gradually added predictors to the base model that included clinical factors only. Next, we conducted backward stepwise selection with an entry level of 0.15 and a stay level of 0.20 to identify the final parsimonious model in the NHS/HPFS cohorts.

Based on the regression coefficients in the final model, we created a risk scoring system for PPCRC. One point was assigned to the predictor with the smallest positive value of the regression coefficient, and points for other predictors were assigned based on the ratio of their corresponding regression coefficients to the minimum positive coefficient. The points were kept to one decimal place and then summed up to generate the PPCRC risk score for each participant.

For the missing data in demographic and lifestyle predictors in the NHS/HPFS cohorts, we leveraged the repeated assessments and carried forward the non-missing values from prior questionnaire cycles. For clinical predictors and in the MGB cohort we used the missing indicator method. To account for missing data, for each individual, we calculated the score based on the non-missing predictors and then standardised the score by dividing it by the sum of the average score across categories among the non-missing predictors.

In the MGB cohort, there were two endoscopy quality variables (quality of bowel preparation [excellent, good, fair, poor] and cecal intubation [yes/no]) that were not available in the NHS/HPFS cohorts but have been shown to be important for PPCRC.[8] Thus, we added these to the final model for validation in the MGB cohort and updated the PPCRC score based on the regression coefficients, using the algorithm described above.

We next applied the final PPCRC score to the MGB cohort for validation by calculating the score for each participant. We classified patients into low- and high-score groups using the median value in the entire cohort as the cutoff. We plotted Kaplan–Meier curves for the cumulative incidence of PPCRC according to the low and high PPCRC scores. To assess the added value of the PPCRC score for tailored surveillance beyond the current USMSTF recommendations, we conducted the Kaplan–Meier curve analysis among patients with the low- and high-risk polyps separately based on the USMSTF classifications (see details above). We then assessed the association between the PPCRC score and risk of PPCRC using Cox proportional hazards regression and calculated the Uno's C-statistics. Finally, we

calculated the Net Reclassification Improvement (NRI) to assess to what extent the PPCRC score can improve classification of patients for their PPCRC risk, again using the USMSTF risk classification as the reference.[43] NRI is an index to quantify how well a new model (the PPCRC score) reclassifies patients–either appropriately or inappropriately–as compared to an old model (the USMSTF risk classification). We calculated NRI among patients with PPCRC as $NRI_{PPCRC} = (\hat{p}_{up,ppcrc} - \hat{p}_{down,ppcrc})$, where $\hat{p}_{up,ppcrc}$ and $\hat{p}_{down,ppcrc}$ represent the percentage of PPCRC cases that moved up and down, respectively, in the risk spectrum according to the new model compared to the old model. NRI for the total cohort was calculated as $NRI_{total} = (\hat{p}_{up,ppcrc} - \hat{p}_{down,ppcrc}) - (\hat{p}_{up,non-ppcrc} - \hat{p}_{down,non-ppcrc})$.

To assess the robustness of our findings, we conducted several sensitivity analyses by (1) excluding the endoscopy quality measures from the model, (2) stratifying by the use of surveillance colonoscopy, (3) excluding participants with incomplete colonoscopy (not reaching the cecum) or with a fair or poor bowel preparation at baseline, (4) excluding participants with BMI<18.5 kg/m^2, (5) examining the PPCRC score in quintiles, (6) assessing the prediction for proximal colon and distal/rectal cancers separately, (7) including diet to the prediction model, (8) using multiple imputation to handle missing data, and (9) stratified analysis by sex. Details of the sensitivity analyses are described in the Supplementary Materials: sensitive analysis (pp. 2).

All the analysis were performed using SAS software, version 9.4 (SAS Institute, Cary, NC, USA) and STATA™ software, version 17.0 (Stata Corp, College Station, Texas, USA). P values < 0.05 was considered statistically significant.

Role of the funding source

The funders had no role in study design, data collection, data analyses, interpretation, or writing of the report.

Results

During a median follow-up of 12.8 years (interquartile range (IQR): 9.3, 16.7) and 5.1 years (IQR: 2.7, 7.8), respectively, we documented 220 and 241 PPCRC cases among 26,741 and 76,603 patients in the NHS/HPFS cohorts and the MGB cohort (Fig. 1 A and B).

Table 1 presents participants' characteristics at the time of index endoscopy. The median age was 59.9 years (standard deviation (SD 9.9)) and 59.5 years (SD 11.4), with the median age of PPCRC diagnosis of 63.4 years (SD 9.9) and 66.0 years (SD 12.1), in the NHS/HPFS cohorts and MGB cohort, respectively. There were 75% of women in the NHS/HPFS cohorts and 50% in the MGB cohort.

The predictors showed largely similar associations with PPCRC in the NHS/HPFS and MGB cohorts (Supplementary Table S1). Adding demographic and lifestyle factors to the clinical-only model significantly improved the C-statistics (Supplementary Table S2).

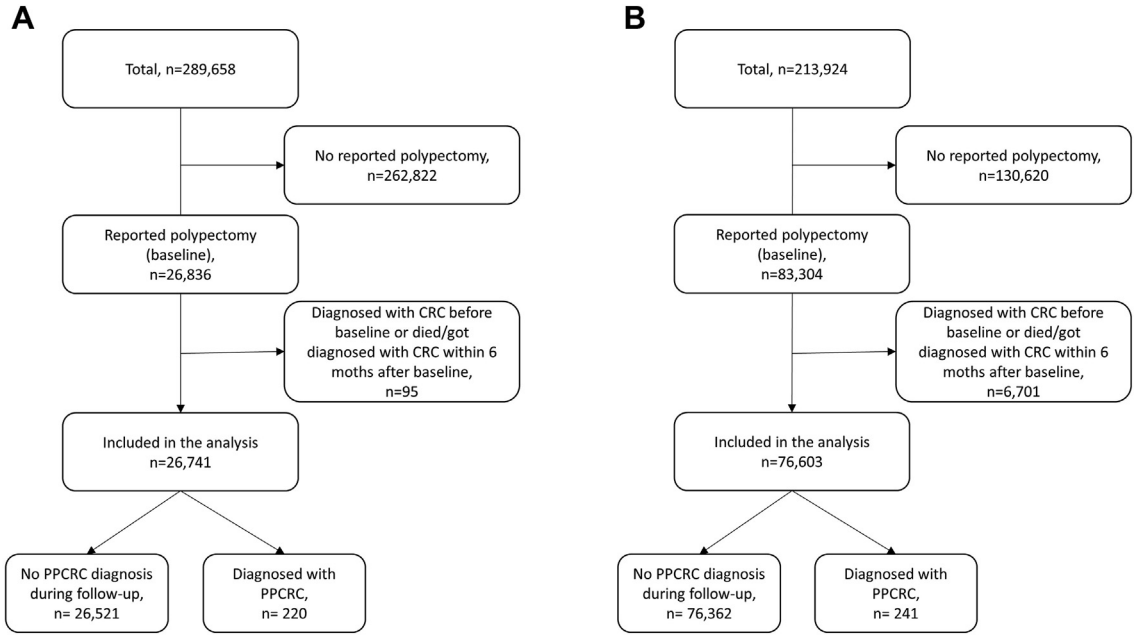


Fig. 1: Flow of study participants. A) the Nurses' Health Study (NHS) I/II the Health Professionals Follow-up Study (HPFS). B) the Mass General Brigham (MGB) Colonoscopy Cohort.

	NHS/HPFS cohorts		MGB Colonoscopy Cohort	
	Overall (n = 26,741)	PPCRC patients (n = 220)	Overall (n = 76,603)	PPCRC patients (n = 241)
Mean age (SD), at index endoscopy, years	59.9 (9.9)	63.4 (9.9)	59.5 (11.4)	66.0 (12.1)
Age groups				
40–44	1324 (5)	9 (4)	5177 (7)	8 (3)
45–49	2811 (11)	10 (5)	4797 (6)	8 (3)
50–54	4711 (18)	20 (9)	17,275 (23)	21 (9)
55–59	4371 (16)	37 (17)	11,321 (15)	32 (13)
60–64	4532 (17)	34 (15)	12,686 (17)	35 (15)
65–69	4057 (15)	45 (20)	10,576 (14)	35 (15)
≥70	4935 (18)	65 (30)	14,771 (19)	102 (42)
Female sex	19,967 (75)	138 (63)	38,173 (50)	112 (46)
Whites	25,410 (95)	206 (94)	60,571 (79)	195 (81)
Family history of colorectal cancer	6565 (25)	64 (29)	11,116 (15)	25 (10)
Histology of index polyps				
Tubular adenomas	11,349 (42)	60 (27)	31,962 (42)	84 (35)
Serrated polyps	9818 (37)	62 (28)	27,307 (36)	65 (27)
Tubulovillous adenomas	2304 (9)	43 (20)	2167 (3)	14 (6)
Villous adenomas	488 (2)	19 (9)	238 (0.3)	2 (1)
CIS/high-grade dysplasia	424 (2)	9 (4)	816 (1)	17 (7)
Having ≥3 adenomas	1494 (6)	23 (10)	11,241 (15)	43 (18)
Having ≥3 serrated polyps	691 (3)	3 (1)	7664 (10)	15 (6)
Large (≥10 mm) adenoma	4697 (18)	74 (34)	7623 (10)	51 (21)
Large (≥10 mm) serrated polyp	911 (3)	5 (2)	2675 (3)	14 (6)
Placement of adenomas				
Proximal colon	9128 (34)	75 (34)	39,845 (52)	144 (60)
Distal colon	8508 (32)	91 (41)	18,609 (24)	62 (26)
Rectum	5382 (20)	69 (31)	7047 (9)	34 (14)
Placement of serrated polyps				
Proximal colon	2830 (11)	20 (9)	13,538 (18)	43 (18)
Distal colon	3883 (15)	16 (7)	14,190 (19)	36 (15)
Rectum	3681 (14)	17 (7)	14,717 (19)	36 (15)
Having ≥1 endoscopies before polypectomy	12,010 (45)	87 (39)	11,848 (15)	30 (12)
Reason for index endoscopy				
Screening	16,811 (63)	117 (53)	58,254 (76)	154 (64)
Symptoms	7442 (28)	70 (32)	13,953 (18)	43 (18)
Lack of cecal intubation in index endoscopy	–	–	5330 (7)	55 (23)
Quality of bowel preparation for index endoscopy	–	–		
Excellent	–	–	15,504 (20)	38 (16)
Good	–	–	44,428 (58)	104 (43)
Fair	–	–	6820 (9)	33 (14)
Poor	–	–	1271 (2)	7 (3)
Body Mass Index, kg/m^2				
<18.5	168 (1)	3 (1)	395 (0.5)	4 (2)
18.5–24.9	10,120 (38)	73 (33)	12,128 (16)	38 (16)
25.0–29.9	9573 (36)	77 (35)	19,734 (26)	77 (32)
≥30.0	6078 (23)	52 (24)	20,381 (27)	64 (27)
Alcohol >1 drink/day for women, >2 drinks/day for men	3811 (14)	42 (20)	–	–
Smoking status			–	–
Never	12,952 (48)	90 (41)	–	–
Past	11,386 (43)	106 (48)	–	–
Current	2282 (9)	21 (10)	–	–
Regular use of aspirin/NSAID ≥2 tablets per week	9924 (37)	76 (35)	–	–
Physical activity ≥7.5 METs-hours per week	19,111 (71)	152 (69)	–	–

Abbreviations: CIS, carcinoma in situ; METs, metabolic equivalent task score; NSAID, non-steroidal anti-inflammatory drugs.

Table 1: **Baseline characteristics of study participants who underwent polypectomy and patients with post-polypectomy colorectal cancer (PPCRC) in the Nurses' Health study (NHS) I/II, Health Professional Follow-up Study (HPFS), and Mass General Brigham (MGB) Colonoscopy Cohort.**

Nine predictors were selected into the final model in the NHS/HPFS cohorts. We further added the two quality measures of index endoscopy available in the MGB cohort that were found to predict PPCRC. Table 2 presents the association of these 11 predictors with risk of PPCRC and their assigned points for the PPCRC score. A higher point indicated a higher risk of PPCRC. Most predictors had a missingness of <15% in all cohorts (Supplementary Table S3).

We applied the PPCRC score to the MGB cohort for validation. The mean score (SD) was 0.3 (0.2). Patients with a high PPCRC score (>0.3 (median)) had a much higher risk of PPCRC than those with a low score (Fig. 2A). Importantly, the same trend was found even within patients with low- and high-risk polyps according to the USMSTF recommendations (Fig. 2B and C). Compared to patients with a low PPCRC score, those with a high score had a more than tripled risk of developing PPCRC among patients with any polyps (HR = 3.55, 95% CI, 2.59, 4.88), with low-risk polyps (HR = 3.05, 95% CI, 2.12, 4.36), and with high-risk polyps (HR = 3.10, 95% CI, 1.50, 6.44) (Table 3). For prediction of PPCRC, the score showed a C-statistic of 0.75 (95% CI, 0.70, 0.79) among patients with any polyps, 0.73 (95% CI, 0.68, 0.78) among low-risk polyps, and 0.71 (95% CI, 0.64, 0.79) among high-risk polyps.

Fig. 3 shows the discrimination and reclassification of PPCRC by the score. Patients who developed PPCRC had a higher score (mean [SD] = 0.5 [0.3]) than the overall MGB cohort. While patients with high-risk polyps had a higher score than those with low-risk polyps (0.5 [0.2] vs. 0.3 [0.2]), the score reclassified 45% (95% CI 36–54%) of the patients with PPCRC based on their predicted risk of PPCRC compared to the classifications used in the current USMSTF recommendations (P < 0.001), where in the total cohort, the NRI was 13% (95% CI, 4–23%, P < 0.006).

We obtained similar results in the sensitivity analyses, as presented in the Supplementary material, Supplementary results (pp. 4–14).

Discussion

To the best of our knowledge, this is the first study to develop a risk prediction model for PPCRC. Using three large prospective cohorts with integrated clinical and epidemiologic data, we developed a risk score for PPCRC based on 11 clinical, demographic, and epidemiologic factors. We further validated the PPCRC score in the large MGB cohort and demonstrated its potential for improved risk stratification among patients with polypectomy. The results provide novel insights into PPCRC and a proof of concept for risk-based colonoscopy surveillance.

The existing clinical guidelines for colonoscopy surveillance, such as that developed by the USMSTF, are

Predictors	Regression coefficient	Hazard ratio (95% confidence interval)	Points assigned[c]
Age groups at index endoscopy, years			
40–44	Ref	Ref	0
45–49	−0.29	0.75 (0.30, 1.87)	−4.6
50–54	−0.02	0.99 (0.43, 2.23)	−0.3
55–59	0.59	1.80 (0.83, 3.93)	9.4
60–64	0.40	1.49 (0.68, 3.29)	6.4
65–69	0.83	2.29 (1.05, 4.99)	13.3
≥70	1.21	3.36 (1.56, 7.23)	19.4
Sex			
Female	Ref.		0
Male	0.24	1.27 (0.90, 1.78)	3.8
Histology of index polyps			
Tubular adenomas	Ref.		0
Serrated polyps	0.13	1.13 (0.72, 1.80)	2.0
Tubulovillous adenomas	0.86	2.37 (1.56, 3.62)	13.8
Villous adenomas	1.46	4.31 (2.49, 7.47)	23.4
CIS/high-grade dysplasia	0.94	2.57 (1.23, 5.37)	15.1
Having ≥3 adenomas			
No	Ref		0
Yes	0.40	1.49 (0.93, 2.37)	6.3
Large (≥10 mm) adenoma			
No	Ref		0
Yes	0.35	1.43 (0.99, 2.06)	5.7
Placement of serrated polyps proximal colon			
No	Ref.		0
Yes	0.50	1.66 (0.99, 2.76)	8.1
Having ≥1 endoscopies before Polypectomy			
Yes	Ref.		0
No	0.19	1.21 (0.91, 1.60)	3.0
Lack of cecal intubation in index endoscopy[a]			
No	Ref.		0
Yes	0.73	2.07 (1.22, 3.53)	11.6
Quality of bowel preparation for index endoscopy[a]			
Excellent	Ref.		0
Good	−0.18	0.84 (0.58, 1.22)	−2.9
Fair	0.27	1.31 (0.82, 2.09)	4.3
Poor	0.36	1.43 (0.63, 3.23)	5.7
Body Mass Index, kg/m^2			
<18.5	1.09	2.96 (0.93, 9.46)	17.4
18.5–24.9	Ref.		0
25.0–29.9	0.06	1.06 (0.77, 1.47)	1.0
≥ 30.0	0.38	1.46 (1.02, 2.09)	6.0
Alcohol >1 drink/day for women, >2 drinks/day for men[b]			
No	Ref.		0
Yes	0.45	1.56 (1.11, 2.20)	7.1

Abbreviation: CIS, carcinoma in situ. [a]These variables were only available in the Mass General Brigham (MGB) Colonoscopy Cohort and thus the Hazard ratios were derived from the MGB cohort. [b]These variables were only available in the NHS/HPFS cohorts. [c]One point was assigned to the predictor with the smallest positive value of the regression coefficient, and points for other predictors were assigned based on the ratio of their corresponding regression coefficients to the minimum positive coefficient. The points were kept to one decimal place and then summed up to generate the PPCRC risk score for each participant.

Table 2: **Association of predictors in the final model for risk of post-polypectomy colorectal cancer (PPCRC) in the Nurses' Health study (NHS) I/II, Health Professional Follow-up Study (HPFS) and the corresponding points used to create the PPCRC score.**

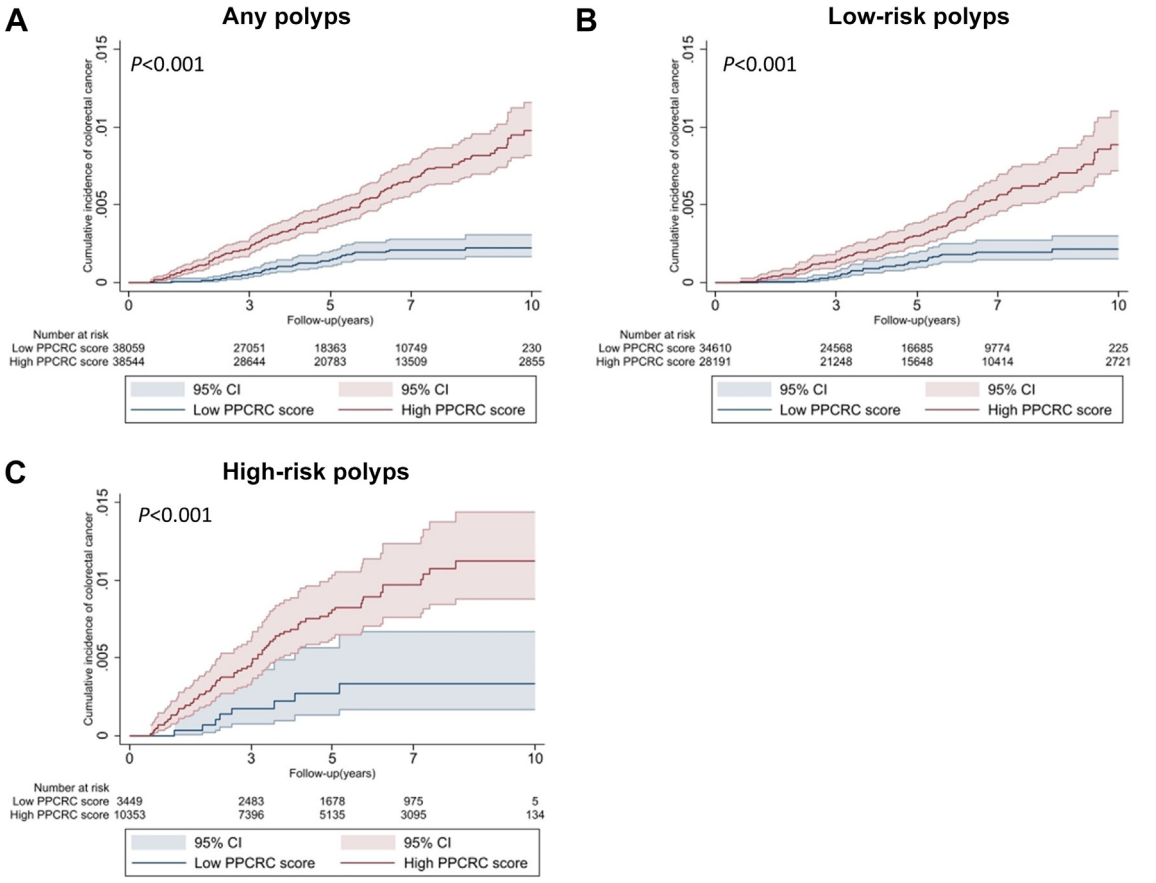


Fig. 2: Kaplan–Meier estimates of cumulative incidence of post-polypectomy colorectal cancer (PPCRC) according to low and high PPCRC score, in the Mass General Brigham (MGB) Colonoscopy Cohort. A) Among patients with any index polyps. B) Among patients with low-risk index polyps* C) Among patients with high-risk index polyps*. *Low- and high-risk polyps were defined based on the 2020 US Multi-Society Task Force recommendations for post-polypectomy colonoscopy surveillance. High-risk polyps were those with a recommended interval for surveillance of 3 years (including 5–10 tubular adenomas <10 mm, adenoma ≥10 mm, adenoma with tubulovillous or villous histology, adenoma with high-grade dysplasia, >10 adenomas, 5–10 sessile serrated polyps <10 mm, sessile serrated polyps ≥10 mm, sessile serrated polyps with dysplasia, hyperplastic polyps ≥10 mm and/or, traditional serrated adenomas as their most severe findings) and low-risk polyps were all other polyps (including patients with 1–2 tubular adenomas <10 mm, ≤20 hyperplastic polyps in the rectum or sigmoid colon <10 mm, ≤20 hyperplastic polyps proximal to sigmoid colon <10 mm, 1–2 sessile serrated polyps <10 mm, 3–4 tubular adenomas <10 mm, or 3–4 sessile serrated polyps <10 mm as their most severe findings). **P-value below 0.05 was considered statistically significant.

solely based on histopathologic features of index polyps.[7] Consistent with prior data, we showed that features of high-risk polyps, such as advanced histology and large size of adenomas and proximal serrated polyps, were associated with higher PPCRC risk[27,44–49] and included in the PPCRC score. However, as shown in Fig. 2, there is a great variation in the risk of CRC among individuals with high-risk polyps, some of whom developed CRC within a year of index colonoscopy while others were free of CRC for more than 10 years. On the other hand, prior data are limited and inconsistent for other polyps, such as non-advanced adenomas, proximal serrated polyps, and multiple HPs, with some studies reporting an increased risk of CRC[29] whereas others observing a similar or even lower risk compared to those with no polyps.[27,44–49] Through comprehensive assessment of index polyps according to histology, size, sublocation, and multiplicity, our study provides novel data regarding the independent effect of these features on PPCRC risk. Furthermore, by developing a numerical algorithm to weigh these features according to their associations with PPCRC, we accounted for the variability in their influence on PPCRC and provided a rigorous assessment of PPCRC risk.

Besides polyp features, colonoscopy quality has been identified as another important predictor for PPCRC.[41] High-quality examinations are essential to reduce the likelihood of missed or incompletely resected lesions at colonoscopy.[50] Indeed, we showed that two major quality measures—quality of bowel preparation and cecal

Risk of index polyps according to the US Multi-Society Task Force	Low PPCRC score (−0.08, 0.30)	High PPCRC score (>0.30, 1.42)	HR per 1-SD PPCRC score increase[b]	P-trend[c]	C-statistic[d]
Any index polyps					0.75 (0.70, 0.79)
Median (IQR)	0.17 (0.08, 0.23)	0.45 (0.38, 0.57)			
No. of participants (%)	38,059 (50)	38,544 (50)			
No. of PPCRC cases (%)	48 (20)	193 (80)			
HR (95% CI)	Ref.	3.55 (2.59, 4.88)	1.94 (1.76, 2.15)	<0.0001	
Low-risk index polyps[e]					0.73 (0.68, 0.78)
Median (IQR)	0.16 (0.08, 0.23)	0.43 (0.36, 0.53)			
No. of participants (%)	34,610 (55)	28,191 (45)			
No. of PPCRC cases (%)	40 (25)	117 (75)			
HR (95% CI)	Ref.	3.05 (2.12, 4.36)	1.83 (1.61, 2.06)	<0.0001	
High-risk index polyps[e]					0.71 (0.64, 0.79)
Median (IQR)	0.20 (0.13, 0.26)	0.54 (0.43, 0.67)			
No. of participants (%)	3449 (25)	10,353 (75)			
No. of PPCRC cases (%)	8 (10)	76 (90)			
HR (95% CI)	Ref.	3.10 (1.50, 6.44)	1.84 (1.54, 2.21)	0.002	

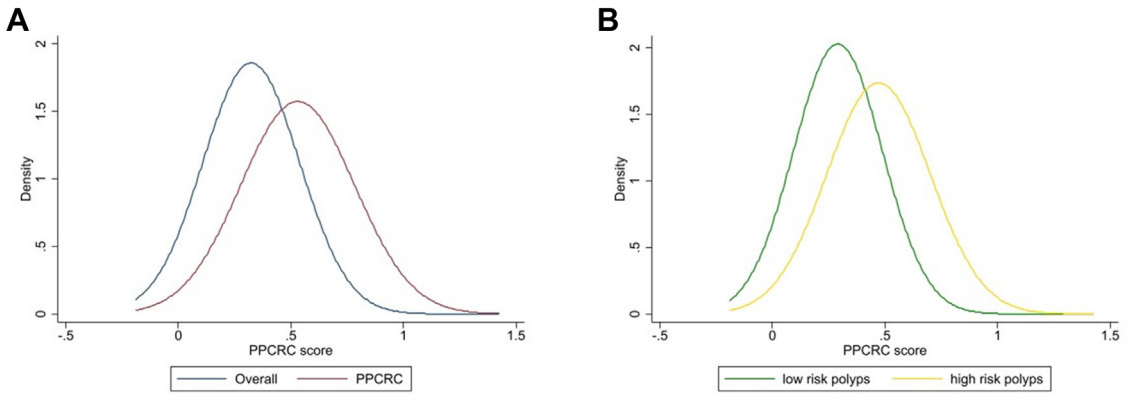
Abbreviations: CI, confidence interval; HR, hazard ratio; IQR, interquartile range; SD, standard deviation. [a]The predictors included histology (tubular adenomas, serrated polyps, tubulovillous adenomas, villous adenomas, carcinoma in situ/high grad dysplasia), having ≥3 adenomas (yes/no), large adenoma (≥10 mm) (yes/no), placement of serrated polyps in the proximal colon), having ≥1 endoscopies before polypectomy (yes/no), quality of bowl preparation for index endoscopy (excellent, good, fair, poor), lack of cecal intubation in index endoscopy (yes/no), age at index endoscopy (40–44, 45–49, 50–54,55–59,60–64,65–69, ≥70), sex (female/male), body mass index (16.5–18.4, 18.5–24.9, 25.0–29.9, ≥30). [b]SD = 0.21 for the PPCRC score. [c]P-value below 0.05 was considered statistically significant. [d]UNO C-statistics for PPCR score as a continuous variable. [e]Low- and high-risk polyps were defined based on the 2020 USMSTF recommendations for post-polypectomy colonoscopy surveillance. High-risk polyps were those with a recommended interval for surveillance of 3 years (including 5–10 tubular adenomas <10 mm, adenoma ≥10 mm, adenoma with tubulovillous or villous histology, adenoma with high-grade dysplasia, >10 adenomas, 5–10 sessile serrated polyps <10 mm, sessile serrated polyps ≥10 mm, sessile serrated polyps with dysplasia, hyperplastic polyps ≥10 mm and/or, traditional serrated adenomas as their most severe findings) and low-risk polyps were all other polyps (including patients with 1–2 tubular adenomas <10 mm, ≤20 hyperplastic polyps in the rectum or sigmoid colon <10 mm, ≤20 hyperplastic polyps proximal to sigmoid colon <10 mm, 1–2 sessile serrated polyps <10 mm, 3–4 tubular adenomas <10 mm, or 3–4 sessile serrated polyps <10 mm as their most severe findings).

***Table 3*: Risk of post-polypectomy colorectal cancer (PPCRC) according to low and high PPCRC score[a] in the Mass General Brigham (MGB) Colonoscopy Cohort.**

intubation—significantly improved the prediction of PPCRC, particularly among patients with high-risk polyps. Consistent with our results, a recent study found that adenoma detection rate (ADR), a metric reflecting colonoscopy quality, was inversely associated with risk of metachronous advanced neoplasia and significantly improved the prediction of metachronous advanced neoplasia compared to the 2020 USMSTF risk classification.[26] Increasing ADR has been recognised as an important target to improve the benefit of CRC screening.[51–53] In our study, we considered the two important determinants of ADR instead of ADR itself because compared to bowel preparation quality and cecal intubation, ADR has not yet been routinely assessed in the clinic and readily available in our EHR system. Nevertheless, our findings support the importance of colonoscopy quality for predicting PPCRC and tailoring surveillance.

In addition to clinical factors, we found that demographic and lifestyle risk factors improved the predictive ability for PPCRC. These results are not unexpected based on the compelling epidemiologic data that older age, male sex, overweight and obesity, and heavy alcohol drinking are established risk factors for CRC.[54] The findings are also consistent with our recent study showing that adherence to a healthy lifestyle after polypectomy was associated with lower risk of PPCRC and all-cause mortality.[28] These data have important clinical implications for prevention of PPCRC and highlight the great value of incorporating other epidemiologic risk factors with index polyp features when making clinical recommendations for colonoscopy surveillance. For example, in the mutually adjusted model we found that some of the risk factors had a similar or even stronger association with PPCRC risk compared to established polyp features that are considered in the current surveillance guidelines (e.g., BMI ≥30.0 kg/m^2 vs. large adenoma; heavy alcohol drinking vs. ≥3 adenomas). Therefore, by accounting for these important risk factors, we may better tailor colonoscopy surveillance to patients who are more likely to benefit and potentially improve clinical practice and cost-effectiveness of CRC prevention.

Our study has several strengths, including the integrated clinical and epidemiologic data, prospective design, large sample size, and long follow-up time of the four cohorts. Furthermore, by using the population-based cohorts for model development and the EHR-based cohort for model validation, we were able to leverage the unique strengths and somewhat complementary data



C

Risk of index polyps according to the US Multi-Society Task Force*	Low PPCRC score**	High PPCRC score**	Total	NRI (95% CI), p value***
Overall cohort				
Low	34,610 (55%)	28,191 (45%)	62,801	
High	3,449 (25%)	10,353 (75%)	13,802	
Total	38,059	38,544	76,603	13% (4-23%), p<0.006
PPCRC cases				
Low	40 (25%)	117 (75%)	157	
High	8 (10%)	76 (90%)	84	
Total	48	193	241	45% (36-54%), p<0.001

Fig. 3: Discrimination and reclassification of patients by the post-polypectomy colorectal cancer (PPCRC) score in the Mass General Brigham (MGB) Colonoscopy Cohort. **A)** Density plot of the PPCRC score among the overall cohort and PPCRC cases. **B)** Density plot of the PPCRC score in participants with low- and high-risk index polyps* **C)** Reclassification by the PPCRC score of participants in the overall cohort and PPCRC cases whose index polyps were classified as low and high risk according to the US Multi-Society Task Force (USMSTF) recommendations and the estimate of the Net Reclassification Improvement (NRI). *Low- and high-risk polyps were defined based on the 2020 USMSTF recommendations for post-polypectomy colonoscopy surveillance. High-risk polyps were those with a recommended interval for surveillance of 3 years (including 5–10 tubular adenomas <10 mm, adenoma ≥10 mm, adenoma with tubulovillous or villous histology, adenoma with high-grade dysplasia, >10 adenomas, 5–10 sessile serrated polyps <10 mm, sessile serrated polyps ≥10 mm, sessile serrated polyps with dysplasia, hyperplastic polyps ≥10 mm and/or, traditional serrated adenomas as their most severe findings) and low-risk polyps were all other polyps (including patients with 1–2 tubular adenomas <10 mm, ≤20 hyperplastic polyps in the rectum or sigmoid colon <10 mm, ≤20 hyperplastic polyps proximal to sigmoid colon <10 mm, 1–2 sessile serrated polyps <10 mm, 3–4 tubular adenomas <10 mm, or 3–4 sessile serrated polyps <10 mm as their most severe findings). **Row percentages are shown after the number of participants. ***P-value below 0.05 was considered statistically significant.

in each of the two study types (e.g., detailed lifestyle vs. clinical data) and to assess the translational potential of our model for guiding colonoscopy surveillance in the clinic.

Some limitations of our study need to be noted. First, because data on quality of bowel preparation and cecal intubation of index endoscopy were unavailable in the NHS/HPFS cohorts, we were not able to validate their contribution to PPCRC prediction. Furthermore, not all lifestyle and diet factors available in the NHS/HPFS were available in the MGB cohort. However, the only predictor in the final model that was missing in the MGB cohort was alcohol. Second, inevitably there were missing data in both the development and validation cohorts. We used missing indicators in the model selection and standardised the PPCRC score to account for missing data in the validation phase. The similar results obtained in the sensitivity analysis using multiple imputation indicate the robustness of our findings. Third, because of the evolving nature and lack of consensus regarding pathologic classification of serrated polyps, we were unable to distinguish HPs from SSPs and TSAs in the NHS/HPFS cohorts. However, expert panels have suggested that a large, serrated polyp is a good indicator of SSPs[35] and we already included size, location, and multiplicity of serrated polyps in the modelling. Fourth, in the NHS/HPFS cohorts, the endoscopy, demographic, and lifestyle factors were all self-reported. Yet, the validity of these self-reported measures has been established in our previous studies.[34–36] Fifth, in the NHS/HPFS cohorts we did not specifically collect surveillance colonoscopy data. Finally,

MGB is a tertiary healthcare system, where patient population and quality of care may be different from community hospitals and the NHS/HPFS cohorts were based on a selected population of health professionals. This, along with the study participants being predominantly white in all our cohorts reduces the generalisability of our findings.

Further studies are needed in several other aspects as well. First, follow-up studies are needed to determine the optimal surveillance intervals for individuals assigned with different PPCRC scores based on our prediction model. Second, it may be helpful to develop a risk calculator based on our model, which will enable clinicians to better communicate with patients regarding their risk and improve the compliance with surveillance colonoscopy. Third, implementation studies are needed to test the feasibility and cost-effectiveness of different strategies of integrating risk assessment tool in clinical practice, which may require development and maintenance of certain infrastructure in the electronic health record system and training of medical staff.[56,57] The risk assessment tool also needs to have good acceptance by both providers and patients.[57] Fourth, it would be interesting to investigate if newer modelling approaches, such as machine learning, may help improve the performance of our developed model. Finally, more studies are needed to assess the added value of other predictors, such as ADR and molecular features of polyps, as these are established risk factors for CRC,[51–53,58] beyond the predictors included in the current study.

We developed and validated a risk stratification model for PPCRC that may be useful to tailor colonoscopy surveillance based on the information readily available in the EHR or easily collected in the clinic. Our results indicate the important of patients' demographic and lifestyle characteristics for risk assessment of PPCRC and recommendations of colonoscopy surveillance.

Contributors

Study concept and design: MS and MDK. Generation, collection, assembly of data: KWu, LW, XH, DH, ATC and MS. Statistical analysis: MDK, ZF, MW, and MS. Interpretation of data: MDK, KWa, LW, GP, PB, ZF, EG, SO, MW, and MS. Drafting of the Manuscript: MDK. Critical revision of the manuscript for important intellectual content: KWa, LW, GP, PB, KWu, XH, DH, ZF, SO, ATC, EG, and MS. Obtained funding: MS, SO, and PB. Study supervision: MS. Author(s) responsible for the decision to submit the manuscript: MDK and MS. MDK, KWa and MS accessed and verified the underlying data. All authors approved the final version.

Data sharing statement

The data for the Nurses' Health Study I and II and Health Professionals Follow-up Study that support the findings of this study are available on request at https://www.nurseshealthstudy.org/researchers (email: nhsaccess@channing.harvard.edu) and https://sites.sph.harvard.edu/hpfs/for-collaborators/. The data for the Mass General Brigham Colonoscopy Cohort that support the findings of this study and further information about the study may be available upon request to the corresponding author and approval by the Institutional Review Board.

Declaration of interests

PB declares support from the South-Eastern Norway Regional Health Authority. SO declares support from the US National Cancer Institute. KWu declares that she is currently an employee and stockholder of Vertex Pharmaceuticals. This study was not funded by this commercial entity. All other authors declare no competing interests.

Acknowledgements

We thank the participants and staff of the Nurses' Health Study and Health Professionals Follow-up Study for their valuable contributions. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centres for Disease Control and Prevention's National Program of Cancer Registries (NPCR) and/or the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centres. Participating central cancer registries include the following American states: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was supported by the US National Institutes of Health (NIH) grants (U01CA261961; R35 CA197735; UM1 CA186107; P01 CA87969; U01 CA176726, U01 CA167552); the American Cancer Society (MRSG-17-220-01 – NEC); the South-Eastern Norway Reginal Health Authority (project number: 20190770); the Deutsche Forschungsgemeinschaft (DFG) (Project number: 426308975).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2023.102139.

References

1 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(1):7–34.
2 Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The national polyp study workgroup. *N Engl J Med*. 1993;329(27):1977–1981.
3 Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med*. 2012;366(8):687–696.
4 Winawer SJ, Zauber AG, Fletcher RH, et al. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US multi-society task force on colorectal cancer and the American cancer society. *CA Cancer J Clin*. 2006;56:143–159.
5 Pignone M, Rich M, Teutsch SM, Berg AO, Lohr KN. Screening for colorectal cancer in adults at average risk: a summary of the evidence for the U.S. preventive services task force. *Ann Intern Med*. 2002;137(2):132–141.
6 Manser CN, Bachmann LM, Brunner J, Hunold F, Bauerfeind P, Marbet UA. Colonoscopy screening markedly reduces the occurrence of colon carcinomas and carcinoma-related death: a closed cohort study. *Gastrointest Endosc*. 2012;76(1):110–117.
7 Gupta S, Lieberman D, Anderson JC, et al. Recommendations for follow-up after colonoscopy and polypectomy: a consensus update by the US multi-society task force on colorectal cancer. *Gastroenterology*. 2020;158(4):1131–1153.e5.
8 Lieberman D, Gupta S. Does colon polyp surveillance improve patient outcomes? *Gastroenterology*. 2020;158(2):436–440.
9 Lieberman DA, Williams JL, Holub JL, et al. Colonoscopy utilization and outcomes 2000 to 2011. *Gastrointest Endosc*. 2014;80(1):133–143.
10 Peery AF, Crockett SD, Murphy CC, et al. Burden and cost of gastrointestinal, liver, and pancreatic diseases in the United States: update 2018. *Gastroenterology*. 2019;156:254–272.
11 Joseph DA, Meester RG, Zauber AG, et al. Colorectal cancer screening: estimated future colonoscopy need and current volume and capacity. *Cancer*. 2016;122(16):2479–2486.
12 Wolf AMD, Fontham ETH, Church TR, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA Cancer J Clin*. 2018;68(4):250–281.

13 Ugai T, Sasamoto N, Lee HY, et al. Is early-onset cancer an emerging global epidemic? Current evidence and future implications. *Nat Rev Clin Oncol*. 2022;19(10):656–673.
14 Weinberg DS, Schoen RE. Preneoplastic colorectal polyps: "I found them and removed them-now what?". *Ann Intern Med*. 2019; 171(9):667–668.
15 Robbins EC, Wooldrage K, Cross AJ. Is surveillance colonoscopy necessary for all patients with bowel polyps? *BMJ*. 2020;369:m1706.
16 Meester RGS, Lansdorp-Vogelaar I, Winawer SJ, Zauber AG, Knudsen AB, Ladabaum U. High-intensity versus low-intensity surveillance for patients with colorectal adenomas: a cost-effectiveness analysis. *Ann Intern Med*. 2019;171(9):612–622.
17 Atkin WS, Valori R, Kuipers EJ, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition–Colonoscopic surveillance following adenoma removal. *Endoscopy*. 2012;44(Suppl 3):151–163.
18 Cancer Council Australia Surveillance Colonoscopy Guidelines Working Party. *Clinical practice guidelines for surveillance colonoscopy*. Sydney: Cancer Council Australia; 2018. Available from: https://wiki.cancer.org.au/australia/Guidelines:Colorectal_cancer/Colonoscopy_surveillance.
19 Rutter MD, East J, Rees CJ, et al. British society of gastroenterology/association of coloproctology of great britain and Ireland/public health england post-polypectomy and post-colorectal cancer resection surveillance guidelines. *Gut*. 2020;69(2):201–223.
20 Hassan C, Antonelli G, Dumonceau JM, et al. Post-polypectomy colonoscopy surveillance: European society of gastrointestinal endoscopy (ESGE) guideline - update 2020. *Endoscopy*. 2020;52(8):687–700.
21 van Heijningen EM, Lansdorp-Vogelaar I, Steyerberg EW, et al. Adherence to surveillance guidelines after removal of colorectal adenomas: a large, community-based study. *Gut*. 2015;64(10):1584–1592.
22 Chubak J, McLerran D, Zheng Y, et al. Receipt of colonoscopy following diagnosis of advanced adenomas: an analysis within integrated healthcare delivery systems. *Cancer Epidemiol Biomarkers Prev*. 2019;28(1):91–98.
23 Kruse GR, Khan SM, Zaslavsky AM, Ayanian JZ, Sequist TD. Overuse of colonoscopy for colorectal cancer screening and surveillance. *J Gen Intern Med*. 2015;30(3):277–283.
24 Knudsen MD, Wang K, Berstad P, Polychronidis G, Vithayathil M, Song M. Use of surveillance colonoscopy among individuals with removal of high-risk polyps according to the US multi-society task force recommendations. *Cancer*. 2023;129(9):1394–1401.
25 Ko CW, Riffle S, Michaels L, et al. Serious complications within 30 days of screening and surveillance colonoscopy are uncommon. *Clin Gastroenterol Hepatol*. 2010;8(2):166–173.
26 Gupta S, Earles A, Bustamante R, et al. Adenoma detection rate and clinical characteristics influence advanced neoplasia risk after colorectal polypectomy. *Clin Gastroenterol Hepatol*. 2022;21(7): 1924–1936.
27 He X, Hang D, Wu K, et al. Long-term risk of colorectal cancer after removal of conventional adenomas and serrated polyps. *Gastroenterology*. 2020;158(4):852–861.
28 Wang L, Knudsen MD, Lo CH, et al. Adherence to a healthy lifestyle in relation to colorectal cancer incidence and all-cause mortality after endoscopic polypectomy: a prospective study in three U. S. cohorts. *Int J Cancer*. 2022;151(9):1523–1534.
29 Li D, Doherty AR, Raju M, et al. Risk stratification for colorectal cancer in individuals with subtypes of serrated polyps. *Gut*. 2022;71:2022–2029.
30 van Toledo D, JEG IJ, Bossuyt PMM, et al. Serrated polyp detection and risk of interval post-colonoscopy colorectal cancer: a population-based study. *Lancet Gastroenterol Hepatol*. 2022;7(8):747–754.
31 Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Physical activity, obesity, and risk for colon cancer and adenoma in men. *Ann Intern Med*. 1995;122(5):327–334.
32 Grodstein F, Martinez ME, Platz EA, et al. Postmenopausal hormone use and risk for colorectal cancer and adenoma. *Ann Intern Med*. 1998;128(9):705–712.
33 Nimptsch K, Malik VS, Fung TT, et al. Dietary patterns during high school and risk of colorectal adenoma in a cohort of middle-aged women. *Int J Cancer*. 2014;134(10):2458–2467.
34 Chasan-Taber S, Rimm EB, Stampfer MJ, et al. Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology*. 1996;7(1):81–86.
35 Giovannucci E, Colditz G, Stampfer MJ, et al. The assessment of alcohol consumption by a simple self-administered questionnaire. *Am J Epidemiol*. 1991;133(8):810–817.
36 Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med*. 2013;369(12):1095–1105.
37 Vithayathil M, Smith S, Goryachev S, Nayor J, Song M. Development of a large colonoscopy-based longitudinal cohort for integrated Research of colorectal cancer: partners colonoscopy cohort. *Dig Dis Sci*. 2021;67(2):473–480.
38 Nalichowski R, Keogh D, Chueh HC, Murphy SN. Calculating the benefits of a research patient data repository. *AMIA Annu Symp Proc*. 2006;2006:1044.
39 Stampfer MJ, Willett WC, Speizer FE, et al. Test of the national death index. *Am J Epidemiol*. 1984;119(5):837–839.
40 Healey MT, Driscoll K, Cooke M, Averbach A, Cherala S, Gershman ST. *Cancer incidence and mortality in Massachusetts 2014-2018: Statewide Report*. Massachusetts: Massachusetts Department of Public Health; 2023.
41 Rutter MD, Beintaris I, Valori R, et al. World endoscopy organization consensus statements on post-colonoscopy and post-imaging colorectal cancer. *Gastroenterology*. 2018;155(3):909–925.
42 Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med*. 2011;30(10):1105–1117.
43 Shi R, Zhang T, Sun H, Hu F. Establishment of clinical prediction model based on the study of risk factors of stroke in patients with type 2 diabetes mellitus. *Front Endocrinol*. 2020;11:559.
44 Click B, Pinsky PF, Hickey T, Doroudi M, Schoen RE. Association of colonoscopy adenoma findings with long-term colorectal cancer incidence. *JAMA*. 2018;319(19):2021–2031.
45 Cottet V, Jooste V, Fournel I, Bouvier AM, Faivre J, Bonithon-Kopp C. Long-term risk of colorectal cancer after adenoma removal: a population-based cohort study. *Gut*. 2012;61(8):1180–1186.
46 Cross AJ, Robbins EC, Pack K, et al. Colorectal cancer risk following polypectomy in a multicentre, retrospective, cohort study: an evaluation of the 2020 UK post-polypectomy surveillance guidelines. *Gut*. 2021;70(12):2307–2320.
47 Wieszczy P, Kaminski MF, Franczyk R, et al. Colorectal cancer incidence and mortality after removal of adenomas during screening colonoscopies. *Gastroenterology*. 2019;158(4):875–883.
48 Loberg M, Kalager M, Holme O, Hoff G, Adami HO, Bretthauer M. Long-term colorectal-cancer mortality after adenoma removal. *N Engl J Med*. 2014;371(9):799–807.
49 Lee JK, Jensen CD, Levin TR, et al. Long-term risk of colorectal cancer and related death after adenoma removal in a large, community-based population. *Gastroenterology*. 2020;158(4):884–894.
50 Rees CJ, Bevan R, Zimmermann-Fraedrich K, et al. Expert opinions and scientific evidence for colonoscopy key performance indicators. *Gut*. 2016;65(12):2045–2060.
51 Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med*. 2014; 370(14):1298–1306.
52 Kaminski MF, Wieszczy P, Rupinski M, et al. Increased rate of adenoma detection associates with reduced risk of colorectal cancer and death. *Gastroenterology*. 2017;153(1):98–105.
53 Schottinger JE, Jensen CD, Ghai NR, et al. Association of physician adenoma detection rates with postcolonoscopy colorectal cancer. *JAMA*. 2022;327(21):2114–2122.
54 World Cancer Research Fund International/American Institute for Cancer Research. Continuous update project expert report 2018. In: *Diet, nutrition, physical activity and colorectal cancer*. 2018.
55 Rex DK, Ahnen DJ, Baron JA, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. *Am J Gastroenterol*. 2012;107(9):1315–1329.
56 Demb J, Gupta S. Realizing the promise of personalized colorectal cancer screening in practice. *J Natl Cancer Inst*. 2021;113(9):1120–1122.
57 Kastrinos F, Kupfer SS, Gupta S. Colorectal cancer risk assessment and precision approaches to screening: brave new World or worlds apart? *Gastroenterology*. 2023;164(5):812–827.
58 van Lanschot MCJ, Carvalho B, Rausch C, et al. Molecular profiling of longitudinally observed small colorectal polyps: a cohort study. *eBioMedicine*. 2019;39:292–300.