

Supplementary materials

From Algorithms to Clinical Utility: A systematic review of risk prediction models for colorectal cancer

Deborah Jael Herrera , Wessel van de Veerdonk, Daiane Seibert , Claudia Gutierrez Ortiz, Moges Boke , Karen Feyen, Nigus Yimer , Allegra Ferrari, and Guido Van Hal

Table of Contents

Appendix 1: PRISMA 2020 Checklist.....	2
Appendix 2. Evidence Matrix.....	5
Figure S1. Evidence matrix on the statistical method, model performance, generalizability, and clinical usability of risk prediction models for CRC	5
Appendix 3: Extended characteristics of included studies	6
Appendix 4: Risk of bias assessment of all 37 studies	31
<i>Table S1.</i> Detailed appraisal and judgements for the risk of bias assessment using PROBAST	31
Appendix 5. Evaluation of the potential clinical utility of all included risk prediction models for colorectal cancer	42
<i>Table S2.</i> Model performance metrics of risk prediction models for colorectal cancer and their projected clinical utility.....	42
Appendix 6. Search strategies	47

Appendix 1: PRISMA 2020 Checklist

Section and Topic	Item #	Checklist Item	Location Where Item Is Reported
Title			
Title	1	Identify the report as a systematic review.	Page 1
Abstract			
Abstract	2	See the PRISMA 2020 for the abstracts' checklist.	Page 1
Introduction			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Pages 1-2
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Pages 2 and 22 (Section 5)
Methods			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 3 (Section 2.1)
Information sources	6	Specify all databases, registers, websites, organizations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page 4 (Section 2.2)
Search strategy	7	Present the full search strategies for all databases, registers, and websites, including any filters and limits used.	Supplementary Materials Appendix 5
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and, if applicable, details of automation tools used in the process.	Page 4 (Section 2.3)
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and, if applicable, details of automation tools used in the process.	Page 4 (Section 2.4)
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g., for all measures, time points, and analyses), and if they were not, specify the methods used to decide which results to collect.	Supplementary Materials Appendix 2
	10b	List and define all other variables for which data were sought (e.g., participant and intervention characteristics, and funding sources). Describe any assumptions made about any missing or unclear information.	Supplementary Materials Appendix 2
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and, if applicable, details of automation tools used in the process.	Page 4 (Section 2.5)
Effect measures	12	Specify for each outcome the effect measure(s) (e.g., risk ratio and mean difference) used in the synthesis or presentation of results.	Pages 4-5 (Section 2.6, 2.7, and 2.8)
Synthesis	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g., tabulating the study intervention characteristics	Pages 5-6

Section and Topic	Item #	Checklist Item	Location Where Item Is Reported
methods		and comparing against the planned groups for each synthesis (Item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Pages 5-6 (Section 2.7-2.9)
	13c	Describe any methods used to tabulate or visually display the results of individual studies and syntheses.	Pages 6-7 (Section 2.8-2.9)
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If a meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Page 7 (Section 2.9)
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g., subgroup analysis and meta-regression).	Not applicable
	13f	Describe any sensitivity analyses conducted to assess the robustness of the synthesized results.	Not applicable
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Pages 4-5 (Section 2.5), Supplementary Materials Appendix 3
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Not applicable
RESULTS			
Study selection	16	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page 7, (Section 3.1. and Figure 2)
Study characteristics	17	Cite each included study and present its characteristics.	Pages 9-12 (Table 2), Supplementary Materials, Appendix 2
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Page 8 (Section 3.2, Figure 3), Supplementary Materials Appendix 3
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimates and its precision (e.g., confidence/credible interval), ideally using structured tables or plots.	Pages 9-12 and 21-24 (Tables 2-3)
Results of syntheses	20a	For each synthesis, briefly summarize the characteristics and risk of bias among contributing studies.	Page 8 (Section 3.2)
	20b	Present results of all statistical syntheses conducted. If meta-analysis was performed, present for each the summary estimate and its precision (e.g., confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Pages 16-18

Section and Topic	Item #	Checklist Item	Location Where Item Is Reported
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Not applicable
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Not applicable
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Pages 8 and 19-25
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Supplementary Materials Appendix 3
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Pages 25-26 (Section 4.1)
	23b	Discuss any limitations of the evidence included in the review.	Pages 26-27 (Section 4.2)
	23c	Discuss any limitations of the review processes used.	Pages 26-27 (Section 4.2)
	23d	Discuss implications of the results for practice, policy, and future research.	Pages 27-29 (Sections 4.2-4.3)
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Page 3 (Section 2)
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page 3 (Section 2)
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Page 3 (Section 2)
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Pages 29-30
Competing interests	26	Declare any competing interests of review authors.	Page 30
Availability of data, codes, and other materials	27	Report which of the following are publicly available and where they can be found template data collection forms: data extracted from included studies, data used for all analyses, analytic codes, and any other materials used in the review.	Page 29

From: Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D., et al. "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews." BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: <http://www.prisma-statement.org/>

Appendix 2. Evidence Matrix

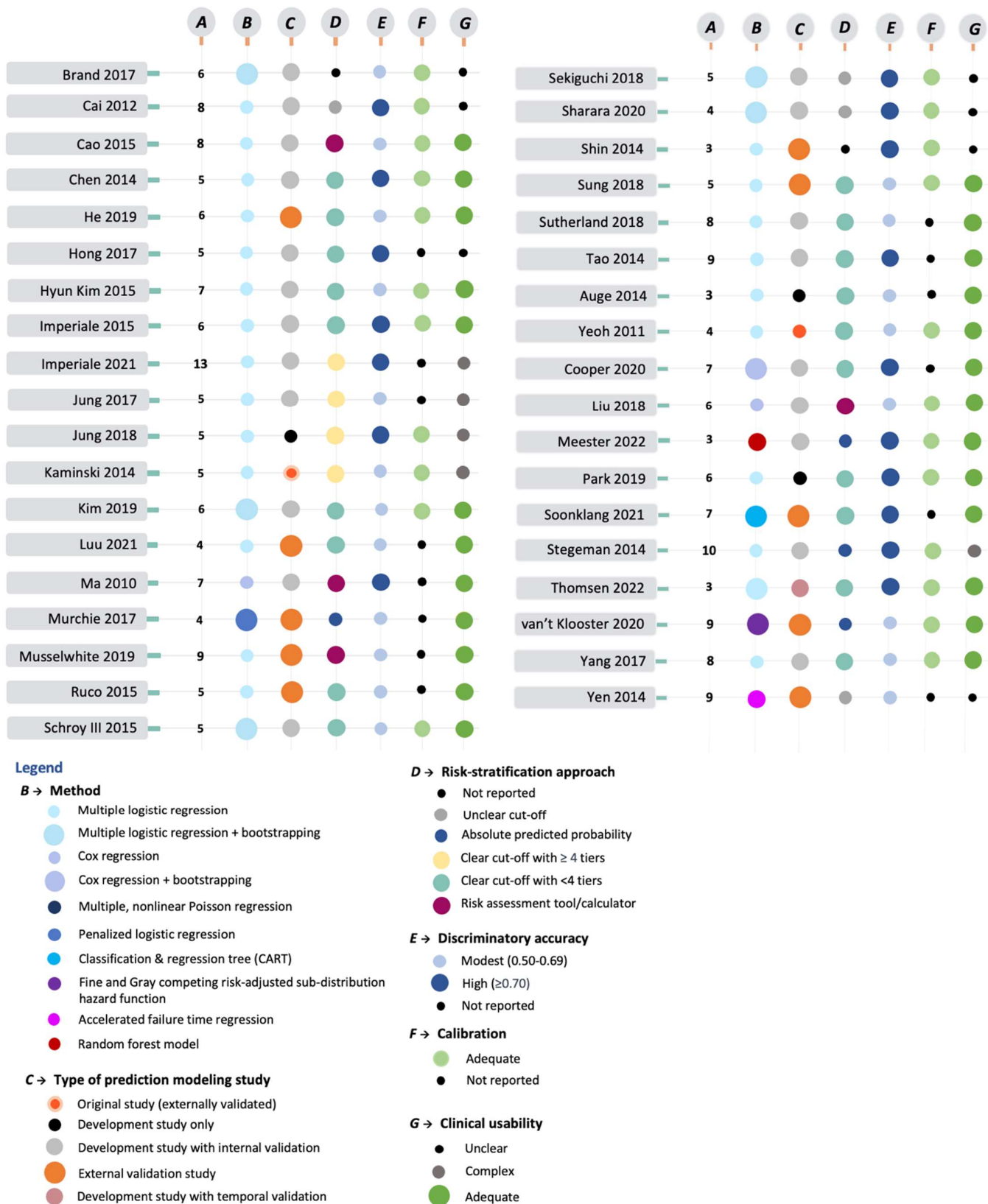


Figure S1. Evidence matrix on the statistical method, model performance, generalizability, and clinical usability of risk prediction models for CRC^{25–53,55–58,60–62,64}. Clinical usability was assessed according to the number of predictors, simplicity of the reported risk-stratification approach, or use of risk assessment tool/calculator for

estimating the risk of CRC. Meanwhile, generalizability was mainly assessed based on whether it has been externally validated.

Appendix 3: Extended characteristics of included studies

Arnau-Collell 2022

Primary objective	To evaluate the potential of using a polygenic risk score (PRS) to improve colorectal cancer screening.
Study location(s)	Barcelona, Spain.
Recruitment period	2009–2019.
Population	(a) Patients aged between 50 and 69 years. (b) 2,893 participants undergoing FIT-test and considering colonoscopy.
Exclusion criteria	Participants with: (1) Two or more first-degree relatives diagnosed with colorectal cancer or adenomatous polyposis. (2) One first-degree relative diagnosed with colorectal cancer or adenomatous polyposis before the age of 60.
Predictor measures	Sex, age, FIT value, and polygenic risk score.
Outcome measures	<i>High-risk lesions such as colorectal cancer, intramucosal carcinoma, high-risk adenomas (HRA), intermediate risk adenoma, and polyposis cases.</i> The histological classification for polyps and cancer followed these guidelines: (1) It adhered to the World Health Organization (WHO)'s criteria. (2) Based on further examination: - Low-risk adenomas (LRAs) were identified. - High-risk adenomas (HRAs) were identified. - All invasive carcinomas were classified as colorectal cancer. (3) Carcinoma in situ, while currently classified as HRA, was also included in the classification.
Statistical approach	Logistic regression (LR) model.
Limitations	(1) Cohort sample size or the number of analyzed genetic variants to calculate PRS may probably not be large enough to reach stronger conclusions. (2) It is also feasible that including additional information in the PRS model, such as gut microbiota data, could enhance risk prediction. (3) Other potential influential predictors were unaccounted for, including body mass index, metabolic syndrome, smoking, anti-inflammatory drugs, and antibiotic use.
New findings	PRS could help improve the current results in colorectal cancer screening programs.

Auge 2014

Primary objective	To investigate whether individuals with positive results from quantitative FITs, in combination with other factors, could be identified as being at greatest risk for advanced colorectal neoplasia.
Study location(s)	Barcelona, Spain.
Recruitment period	July 2011 to December 2012.
Population	(a) Asymptomatic, FIT-positive individuals who participated in the first round of Barcelona CRC screening program, aged 50–69 years. (b) 3109 subjects out 197, 839 were FIT-positive.
Exclusion criteria	Participants were permanently excluded if they had the following: (1) A personal history of colorectal cancer, adenoma, or inflammatory bowel disease. (2) A family history of hereditary or familial colorectal cancer, specifically if they had two first-degree relatives with colorectal cancer or one relative diagnosed before 60 years of age. (3) A severe coexisting illness. (4) Undergone a total colectomy within the past 5 years. Temporary exclusions applied to participants who (6) had tests for fecal blood within the past 2 years, (7) underwent sigmoidoscopy or colonoscopy within the last 5 years, and (8) displayed symptoms necessitating further examination. <i>Note: Those who had prior screening tests could participate once enough time had passed since their last test. Similarly, those displaying symptoms could join if subsequent clinical evaluations turned out to be negative.</i>

Predictor measures	Established risk factors, including age and sex combined with FHbC test results.
Outcome measures	<p><i>Patients with large serrated polyps (≥10 mm) or those with dysplasia were classified as advanced adenomas, whereas patients with serrated polyps less than 10 mm with no dysplasia were classified as nonadvanced adenomas.</i></p> <p>Participants who tested positive with a FIT result of 20 mg/g or more underwent a colonoscopy. This procedure was carried out by the following:</p> <ul style="list-style-type: none"> - Expert endoscopists with experience in conducting more than 400 colonoscopies annually. <p>The detected lesions were then assessed:</p> <ul style="list-style-type: none"> - By specialist pathologists who focus on gastrointestinal oncology. - Using the European guidelines for ensuring quality in colorectal cancer screening and diagnosis. <p>Serrated lesions were categorized based on:</p> <ul style="list-style-type: none"> - The recommendations of the European Society of Gastrointestinal Endoscopy. - Criteria that took into account both the lesion size and the presence of dysplasia.
Statistical approach	Multiple logistic regression model with bootstrapping for the internal validation of the model.
Limitations	<p>(1) Inclusion of only FIT-positive individuals and predominantly comprised members of the Caucasian population, thus limiting the generalizability of the prediction model.</p> <p>(2) Potential confounding bias due to unobserved established risk factors, including BMI, toxic habits, and diet.</p> <p>Limitation identified by reviewers:</p> <ul style="list-style-type: none"> - Low sample size; - Unclear internal validation approach (presented the calibration of the model, but it was unclear how it was estimated or obtained).
New findings	For individuals with positive FIT results: Fecal hemoglobin concentration (FHbC), alongside sex and age, can be employed to determine the likelihood of detecting advanced colorectal neoplasia. Utilizing these factors can help prioritize high-risk individuals to undergo a colonoscopy examination.

Brand 2017

Primary objective	To develop and validate a prediction model for adenoma detection in an effort to determine if physicians' adenoma detection rates (ADRs) should be adjusted for patient-related factors
Study location(s)	United States—the Endoscopic Quality Improvement Program-3 (EQUIP-3) study conducted in nine centers within the country.
Recruitment period	September 2013 to January 2015.
Exclusion criteria	<p>Patients under 50 years of age were excluded if they had a known increased risk of colorectal neoplasia, which typically was not considered in ADR calculations. This includes the following individuals:</p> <ol style="list-style-type: none"> (1) Those who had a colonoscopy due to a high-risk genetic colorectal cancer (CRC) syndrome. (2) Those who were under surveillance colonoscopy because of inflammatory bowel disease. (3) Those who had a personal history of CRC, as these patients likely underwent colorectal surgery.
Predictor measures	Possible precolonoscopic predictors for adenoma detection were selected based on previously published prediction models for the detection of adenomas. All information was obtained from the GIQuIC form by a physician or a nurse. Predictor assessment was blinded by the outcome.
Outcome measures	<i>The detection of ≥ 1 histologically confirmed colorectal adenoma per patient.</i> Histological assessment was performed in daily practice, and, thus, it was not completely blinded for patient factors such as sex and age. However, they were blinded by the BMI, ASA class, race, and ethnicity predictors.
Statistical approach	Multivariable logistic regression.
Limitations	Potential recall bias due to self-reported measures of predictors.
New findings	Age, sex, BMI, ASA class, the distinction between surveillance and screening, and Hispanic or Latino ethnicity were identified as key patient-related predictors for detecting colorectal adenoma. The significant differences in Adenoma Detection Rates (ADRs) are not solely attributable to patient-related factors. Other aspects, possibly involving the physician or technical matters, might contribute to this variation.

Briggs 2022

Primary objective	To evaluate the benefit of combining polygenic risk scores with the QCancer-10 (colorectal cancer) prediction model for non-genetic risk to identify people at highest risk of colorectal cancer.
Study location(s)	United Kingdom.
Recruitment period	March 2006 to July 2010.
Population	UK Biobank participants. Training cohort: n= 30,000 (446 cases). Test cohort: n= 280 664 (4230 cases). For the integrated modeling cohorts: Male integrated modeling cohort: n= 196 091 (1985 cases); and female integrated modeling cohort: n= 238 496 (1458 cases).
Exclusion criteria	Participants were excluded if their information have been redacted by UK Biobank study, have withdrawn their consent, have sex chromosome aneuploidy, have excess relatives, or have sex mismatch. Cases with a diagnosis of colorectal cancer before cohort entry were also excluded.
Predictor measures	For the polygenic risk scores, the summary data from 14 colorectal cancer GWAS cohorts were meta-analyzed to determine the effect sizes of single nucleotide polymorphisms. For the Qcancer-10 predictors, ethnic group, previous medical history, alcohol use and smoking status, and family history of CRC were all obtained from self-reported data in baseline touch-screen responses and verbal interviews at UK Biobank assessment centers.
Outcome measures	The primary outcome in all models was colorectal diagnosis, identified through self-reported at UK Biobank study enrolment visit and International Classification of Disease-9 (153, 154.0, 154.1) and International Classification of Disease-10 (C18-C20) codes in linked cancer and death registries and hospital data.
Statistical approach	Cox proportional hazard.
Limitations	(1) The UK Biobank population differs from the general population in terms of health and demographics. (2) Follow-up data were limited to a median of seven years. (3) Model performance was less robust in women. (4) The Qcancer-10 model may perform differently due to age distribution differences. (5) Lack of representation of mendelian colorectal cancer syndromes in the genetic model. (6) Limited data on colorectal polyps and precursors. (7) PRS models are predominantly developed for individuals of European ethnicity and may not transfer accurately to other populations, leading to potential health inequalities.
New findings	When combined, polygenic risk scores and Qcancer-10 slightly improve risk prediction compared to just using Qcancer-10. But since Qcancer-10 data are easily accessible from health records, the justification for using polygenic risk scores in colorectal cancer population screening is unclear. Before widespread adoption, a deeper evaluation of the benefits, cost-effectiveness, and acceptability of polygenic risk scores in actual screening scenarios is needed.

Cai 2011

Primary objective	To develop and validate a new clinical prediction rule to stratify risk for advanced neoplasms in average-risk individuals.
Study location(s)	19 participating hospitals in China.
Recruitment period	July 1, 2006, to December 31, 2008.
Exclusion criteria	A positive fecal occult blood test within 6 months before referral; iron-deficiency anemia within 6 months before referral, rectal bleeding or hematochezia within the preceding 12 months; an unintentional weight loss of more than 4.5 kg within the preceding 6 months; a recent marked change in bowel habits; a history of colorectal cancer, adenomas, inflammatory bowel disease, or a hereditary polyposis syndromes; prior colonic surgery; a colonic examination within the previous 5 years; a family history of cancer of any type; a medical condition that could increase the risk associated with colonoscopy; and pregnancy.
Predictor measures	A self-administered questionnaire was used before colonoscopy to obtain the potential predictors.
Outcome measures	Colonoscopy was performed by expert endoscopist to detect advanced neoplasms. <i>Colorectal neoplasm: defined as invasive cancer, an adenoma of 10mm or more, a villous adenoma, or an adenoma with high-grade dysplasia.</i>

Statistical approach	Multiple logistic regression for deriving the prediction rule and bootstrapping to internally validate the derived model.
Limitations	<ol style="list-style-type: none"> (1) External validation might not represent the general population, as participants were likely skewed towards being more health-conscious. (2) The Chinese healthcare and insurance systems might have favored participants of higher socioeconomic status, limiting the study's generalizability to disadvantaged groups. (3) Absence of a cost-benefit analysis.
New findings	A new prediction rule comprising age, sex, smoking, diabetes mellitus, green vegetables, pickled food, fried food, and white meat was developed and externally validated, with good discriminatory accuracy and calibration.

Cao 2012

Primary objective	To develop a risk assessment tool for high-risk colorectal adenoma (advanced adenoma or ≥ 3 adenomas) that can be implemented in clinical/general settings through evaluating a comprehensive list of risk factors among White men and women who underwent colonoscopy as their first routine screening for colorectal cancer.
Study location(s)	United States.
Recruitment period	1986 to 2008 for Nurses' Health Study cohort, and 1988 to 2008 for Health Professionals Follow-Up Study cohort.
Exclusion criteria	Individuals who reported occult or visible blood, diarrhea, constipation, or abdominal pain as the indications for the first colonoscopy were excluded. Participants with missing information after carrying forward during predictor assessment were excluded.
Predictor measures	Use of mailed questionnaires at enrollment and 2 years thereafter to collect data on demographics, lifestyle factors, medical history, and disease outcomes, and every 4 years to report update in dietary intake.
Outcome measures	Participants were asked whether polyps had been diagnosed in the past two years, and when a diagnosis was reported, informed consent was obtained to acquire medical records and pathology reports. <i>Men and women diagnosed with at least one adenoma ≥ 1 cm in diameter, or with advanced histology (tubulovillous or villous histologic features or high grade or severe dysplasia), or \geq adenoma at their first colonoscopy were considered as high-risk adenomas and cases in the analysis.</i>
Statistical approach	Stepwise multiple logistic regression and used a ten-fold cross-validation to internally validate the models.
Limitations	<ol style="list-style-type: none"> (1) The absolute risk of high-risk adenoma observed (96.7% in men and 3.8% in women) was comparable to other studies. However, were limited in power to fully assess the impact of extreme risk factors because health professionals are slightly more health conscious than the general US population. (2) The study was only applicable to White men and women, and colonoscopy or sigmoidoscopy was self-reported. However, misclassification was unlikely, as colonoscopy requires sedation.
New findings	A separate risk prediction model for US men and women was developed to predict high-risk adenoma.

Chen 2014

Primary objective	To establish the risk scoring system towards the advanced colorectal neoplasm (ACN) risk in the average-risk populations in the southern Jiangsu Province and to evaluate its effectiveness as the screening tool ACN.
Study location(s)	Southern Jiangsu Province, China.
Recruitment period	July 2011 to December 2012.
Population	<ol style="list-style-type: none"> (a) Asymptomatic, Han nationality, aged ≥ 40 years old. (b) 905 subjects, 48 cases with advanced CN were involved in the study.
Exclusion criteria	The first- and second-degree relatives have a CN history; under-60-years-old first-degree relative has a history of AP or familial hereditary syndromes (including familial AP, hereditary non-polyposis CN, Turcot syndrome, old field syndrome, etc.) and a disease history of CN or polypus disease, inflammatory bowel disease, other organ tumors, etc. Had iron deficiency anemia or the fecal occult blood test was positive, hematochezia, significant weight loss, tenesmus, and other symptoms; performed the colonoscopy in the near five years; or had the colorectal surgery history.

Predictor measures	Self-administered questionnaire covering the demographic characteristics, past medical history, surgical history, medication history, smoking history, alcohol drinking history, tea drinking history, physical activity, diet habits, and defecation frequency.
Outcome measures	The colonoscopy was performed by two experienced by gastrointestinal-endoscopy experts to detect ACN is the outcome of the study. The ACN includes both advanced adenoma and invasive carcinoma. <i>The advanced adenomas include adenomas with diameter ≥ 1 cm, villous adenomas (the villous component should at least be 25%) and tubular adenoma while the invasive carcinoma refers to the tumor of which the malignant cells invaded over the muscularis mucosa.</i>
Statistical approach	Multiple logistic regression model with bootstrapping for the internal validation of the model.
Limitations	(1) The sample size was relatively small; therefore, the results of this study still need to be further verified with the larger external populations. (2) A lot of potential risk factors in this cross-sectional study originated from the patients' memories, which might have some recall bias. Limitation identified by reviewers: <ul style="list-style-type: none"> - Dichotomization of age predictor during model derivation. - Unclear justification for using the 2.5 cutoff for the risk stratification.
New findings	(1) The risk scoring system of advanced CN established in this study had good prediction consistency and distinguishing ability, higher sensitivity, and negative predictive value. (2) It can also detect CN or pre-malignant lesions earlier.

Cooper 2020

Primary objective	To determine the availability of data for key predictors and whether this information could be used to inform more accurate screening referral decisions.
Study location(s)	United Kingdom.
Recruitment period	13th May 2009 to 17th January 2017.
Population	(1) English Bowel Cancer Screening Programme (BCSP) derives participant information were used for this study. (2) For stable predictions, it has been recommended that multivariable models include at least 10 outcome events per degree of freedom: <ul style="list-style-type: none"> i. The dataset for multivariable modeling analysis had 1676 CRC and polyp diagnoses and considered 17 degrees of freedom giving 98.59 outcomes per degree of freedom. ii. The dataset for the model with negative FOBTs included only 735 outcome events and considered 16 degrees of freedom, giving 45.94 outcomes per degree of freedom.
Exclusion criteria	Patients were excluded if they had a previous CRC diagnosis or if they had a high-risk condition (hereditary nonpolyposis colorectal cancer, HNPCC) or familial adenomatous polyposis (FAP).
Predictor measures	Age, sex, previous positive FOBT results, previous negative FOBT results, smoking status, alcohol consumption and family history, BMI, hemoglobin, ferritin, MCV, diabetes, and family history of gastrointestinal cancer
Outcome measures	<i>The index date used for survival analysis was the date of the latest BCSP FOBT result. The outcome was a diagnosis of CRC/polyps up to 2 years after the index date (latest FOBT) recorded in a patient's record.</i> <i>On the diagnosis and definition of CRC or polyps, where detailed information is not given.</i>
Statistical approach	Cox regression and multivariable fractional polynomials with backward elimination. Multivariable fractional polynomials (MFPs) allow nonlinear relationships with continuous predictors to be modeled.
Limitations	(1) Missing data can lead to bias in parameter estimates and reduce sample size and generalizability. (2) The use of real-world data in this dataset introduces differential verification of cancer, impacting the evaluation of predictive variables. Participants with positive FOBT results are more likely to undergo colonoscopy and receive faster diagnoses, while those with negative FOBT results rely on follow-up, leading to ascertainment bias. Consequently, the model may overestimate the predictive power of FOBT and other variables used in the current referral pathway, while underestimating the predictive power of variables not included in the pathway. This limitation arises from the utilization of routine data. (3) The data lack specific details on the various diagnostic types employed in a secondary care setting, leading to potential verification bias.
New findings	<ul style="list-style-type: none"> ▪ This research has identified several potential predictors for CRC in a screening population by exploiting the interface between the screening database and primary care records.

- Additional data could be drawn onto the screening database to contribute to a referral algorithm to improve colonoscopy use and to benefit those at highest risk of CRC.

Deng 2023

Primary objective	To develop noninvasive predictive models for EO-CRC and investigate its risk factors in the Chinese population.
Study location(s)	Shanghai Cancer Center (Fudan [FD] cohort) and Shanghai Renji, Hospital (Renji cohort), China.
Recruitment period	January 1, 2015, and November 30, 2021.
Population	1756 participants aged <50 years (RJ cohort, n = 1359; FD cohort, n = 397). EO-CRC and healthy control (HC) groups included 811 patients and 945 colonoscopy-negative participants (young healthy control [y-HC] group) aged <50 years, respectively.
Exclusion criteria	Not reported.
Predictor measures	They included variables that increased risk, such as family history of CRC, smoking, alcohol consumption, processed meat intake, and sweet and fried food intake. Also included were variables that decreased the risk, such as higher education; eggs and coffee intake; and dietary fiber, calcium, and vitamin supplementation.
Outcome measures	Early-onset colorectal cancer.
Statistical approach	Multivariable logistic regression analysis, eXtreme Gradient Boosting (XGBoost), and random forest (RF).
Limitations	<ol style="list-style-type: none"> (1) Retrospective Design and Recall Bias: The study relied on retrospective data, particularly self-reported food and lifestyle behaviors collected after a tumor diagnosis, making it prone to recall bias. (2) Control of Recall Bias: Although the same interviewers completed questionnaires for both cases and controls, and controls with dietary habit changes were excluded (potentially limiting recall bias), this does not entirely eliminate the possibility of such bias. (3) Potential Unaccounted Confounders: Despite accounting for many potential confounding factors in the analysis, there might still be unidentified or unconsidered confounders that could influence the study's results. (4) Lack of Serving Size Data: The Food Frequency Questionnaire (FFQ) used lacked details on serving sizes, which could have an impact on the accuracy and depth of dietary assessments.
New findings	Identified an association between sweet and fried foods and EO-CRC.

He 2019

Primary objective	To improve the existing APCS score and validate the modified model to confirm its effectiveness in screening high-risk groups of CRC among Chinese asymptomatic population and further test the validity of the score combined with FIT.
Study location(s)	China.
Recruitment period	September 2016 to December 2017.
Exclusion criteria	Those with medical history of colorectal cancer, colorectal polyps, or inflammatory bowel disease; or with colonoscopy contraindications were excluded.
Predictor measures	Used risk factors questionnaire.
Outcome measures	Standardized colonoscopy was conducted by experienced endoscopists in a double-blinded approach at all study sites. <i>Advanced colorectal neoplasia (ACN) was defined as CRC or advanced adenoma. Advanced adenoma was defined as adenomas ≥ 10 mm in diameter, villous histological features (at least 25% villous), high-grade dysplasia, or any combination thereof.</i>
Statistical approach	Multivariate logistic regression.
Limitations	<ol style="list-style-type: none"> (1) Predictors were measured through self-reported survey questionnaire. Misestimation of some risk factors is possible. (2) Though the whole screening score showed good discrimination in triaging high-risk population for CAN, the C-statistic for the modified risk score was only slightly higher than that for the original APCS system, which could be attributed to confounding factors that lowered the efficacy.

(3) Selection bias of receiving FIT in the validation may be observed (only 742 out of 1201 participants received FIT).

New findings	Compared to the original score, they introduced diabetes as a key risk factor with 2 points, alcohol consumption with 1 point, and BMI with 1 point. Sex and other targeted risk factors were left out of the modified score because sex showed significance only for colorectal neoplasia but not for ACN.
---------------------	---

Hong 2017

Primary objective	To develop and validate a prediction model for assessing the probability of ACN using a clinical data warehouse.
Study location(s)	Samsung Medical Center, Seoul, Republic of Korea.
Recruitment period	January 2003 to December 2012.
Population	Training set included 24,725 individuals, and validation set included 24,725 individuals.
Exclusion criteria	Participants with incomplete colonoscopy, poor and inadequate bowel preparation, incomplete colonoscopy report about the number and size related to CRN, incomplete pathology report about the histology and dysplasia grade related to CRN, history of previous colonoscopy, history of colorectal polyps, cancer or surgery, and inflammatory bowel disease.
Predictor measures	A self-administered questionnaire was used to identify current smoking status, alcohol drinking frequency, physical activity, family history of colon cancer, history of colorectal polyps/cancer, comorbidities, and regular use of aspirin.
Outcome measures	(a) (How the outcome was measured). (b) <i>Advanced CRN was defined as a cancer or adenoma that was at least 10 mm in diameter and had high-grade dysplasia, villous or tubulovillous histological characteristics, or any combination thereof.</i> For patients with multiple neoplasms, the size and appearance of the neoplasms with advanced pathology or of the largest polyp were reported.
Statistical approach	Multivariable logistic regression.
Limitations	(1) Lack of External Validation: This absence raises question' regarding the model's potential overfitting and its broader applicability. (2) Exclusion of Specific Populations: The exclusion of those unwilling to undergo screening colonoscopy casts doubt on the model's relevance to individuals who cannot or will not have this procedure. (3) Younger Population Studied: The participants included were younger than the typical age for routine screening colonoscopies. This could result in an underestimation of advanced colorectal neoplasia (ACN) prevalence.
New findings	Once externally validated, the model can serve as a clinically useful tool for facilitating shared decision making related to selecting the screening modalities for early detection and prevention of CRC, especially when provider and patient preference differ.

Hyun Kim 2015

Primary objective	To develop and validate a risk stratification-based screening model for predicting advanced colorectal neoplasia in Korea.
Study location(s)	South Korea.
Recruitment period	2006 to 2009.
Population	Asymptomatic person 30–77 years. - The sample size estimation for the validation cohort was based on methodology of adjusting Asia-Pacific Colorectal Screening (APCS) development (by assumption, power 80%, p -value =0.05, OR=2, prevalence of advanced neoplasia 5%. Finally, a 1300 sample size was attained.
Exclusion criteria	Age < 30 and above 75, prior coloscopy, barium enema, poor bowel preparation, history of colorectal cancer and surgery, colorectal cancer related sign and symptoms, and regular use of anti-inflammatory drug.
Predictor measures	Smoking habits, alcohol consumption, past medical history, and family history of CRC in a first-degree relative were determined by interviews conducted by well-trained nurses. A patient was defined as a current smoker if he/she consumed at least 1 pack per week; consumption of any amount of alcohol exceeding 140 g per week was considered a positive history of alcohol use. Diabetes mellitus (DM) was

	defined as a fasting glucose level of ≥ 126 mg/dL or use of hypoglycemic agents and/or insulin. Height and body weight, which were used to calculate BMI, were routinely measured by trained nurses
Outcome measures	All examinations were performed using a standard video colonoscope. All detected polyps were biopsied, except for multiple hyperplastic polyps in the rectosigmoid colon showing. <i>Colorectal adenoma was defined as an adenoma, irrespective of its grade or villous components, whereas advanced neoplasia was defined as a colorectal carcinoma or advanced adenoma (diameter of ≥ 10 mm, high grade dysplasia, or $\geq 25\%$ villous features).</i>
Statistical approach	Logistic regressions model.
Limitations	<ol style="list-style-type: none"> (1) Single-Center Scope: The study was conducted at only one center, which might not represent broader demographics or practices. (2) Participant Bias: The participants tended to be health-conscious and economically well-off, which could affect the generalizability of the findings. (3) Retrospective Design: The study's retrospective nature restricted the range and depth of data that could be gathered. (4) Suboptimal Model Performance: The diagnostic capability of the risk score model was not ideal. When the model was created, there were no established criteria for CRC screening scores to predict advanced neoplasia. (5) Country-Specific Interpretations: The model's diagnostic performance might vary by country, reflecting different populations and healthcare practices. (6) Inclusion of Younger Asymptomatic Subjects: The study incorporated data from asymptomatic individuals under 50, an age group not typically recommended for CRC screening. This was performed to assess the feasibility of a risk stratification model for this younger group.
New findings	The KCS score is a clinically simple and useful parameter for predicting advanced neoplasia in asymptomatic Korean patients. However, racial disparity should be considered in risk stratification-based screening in individual countries.

Imperiale 2015

Primary objective	To create a risk index for AN using the most common risk factors for colorectal neoplasia.
Study location(s)	Midwest central Indiana, United States.
Recruitment period	December 2004 to September 2011.
Population	Patients aged 50 to 80 years undergo initial screening colonoscopy.
Exclusion criteria	Persons with inflammatory bowel disease, those with high-risk family history (polyposis or nonpolyposis colorectal cancer syndrome), and those reporting a history of polyps that required follow-up colonoscopy.
Predictor measures	A self-administered questionnaire was used to obtain candidate risk factors, including sociodemographic and lifestyle factors, before the screening colonoscopy.
Outcome measures	Colonoscopy reports were reviewed and coded by trained personnel who were blinded to survey information. The most advanced findings in the colorectum were coded for both proximal and distal segments. Colonoscopy with no pathology report was assumed to show no neoplasia, if colonoscopy report did not specify that a tissue specimen had been obtained (biopsy or polypectomy). <i>No outcome definition provided.</i>
Statistical approach	Multivariate logistic regression using split sampling for internal validation of the derived model.
Limitations	<ol style="list-style-type: none"> (1) Predetermined Predictors: The predictor variables were chosen based on the existing literature, which may not account for all relevant factors. (2) Imperfect Discrimination: The prediction equation might categorize some patients with advanced neoplasia, including those with colorectal cancer, as low risk. Moreover, while the model gauges the prevalence of advanced neoplasia in the entire colorectum, it does not differentiate between proximal or distal disease. As a result, it cannot suggest the most appropriate and least invasive test. (3) Validation Method: The split-sample method was used for model validation. While this method can check for overfitting, it does not confirm how well the model might perform with entirely different cohorts. (4) Limited Demographics: The study mainly involved White participants, so its applicability to more racially diverse groups is unclear.

New findings	A 5-variable risk index was developed that could help decision making about colorectal cancer screening for persons currently considered to be at average risk, for whom several test options are equally strongly recommended.
---------------------	---

Imperiale 2021

Primary objective	To incorporate additional risk factors of advanced neoplasia (AN) using the same population of average-risk person undergoing first-time screening colonoscopy to derive and test a risk prediction model for AN.
Study location(s)	Indiana University Medical Center in Indianapolis, Indiana.
Recruitment period	December 2004 to September 2011.
Exclusion criteria	Individuals with inflammatory bowel disease, with a high-risk family history (known polyposis or lynch syndrome, >1 first-degree relative with CRC), and those reporting a history of polyps that required surveillance colonoscopy.
Predictor measures	A 50-item survey of sociodemographic features, family history, personal medical history, lifestyle habits, and medication use. Self-reported and self-measured height, weight, and waist using a 72-inch tape measure.
Outcome measures	Advanced neoplasia and adenomas were detected using colonoscopy (performed in standard fashion based on each site's protocol). <i>AN is defined as CRC or advanced precancerous polyp. Advanced precancerous polyps included adenomas or serrated polyps ≥ 1cm or one with villous histology or high-grade dysplasia.</i>
Statistical approach	Multivariate logistic regression model.
Limitations	<ol style="list-style-type: none"> (1) The model's derivation on a predominantly White cohort undergoing a first screening colonoscopy. (2) The inclusion of 13 variables may also be hard for users to both understand and respond accurately. (3) May need an automatic computation to estimate the risk. (4) Finally, there is still a need for external validation of a completely independent cohort to determine the robustness and generalizability of the model.
New findings	A combination of sociodemographic, physical, and lifestyle features provided good discrimination for risk of advanced neoplasia. Their model identified sizeable lower-risk and higher-risk groups (among average-risk persons) for which clinical decision making can be personalized to suggest noninvasive screening and colonoscopy, respectively.

Jung 2017

Primary objective	To establish a risk-stratification model for ACRN in persons aged < 50 years.
Study location(s)	Kangbuk Samsung Hospital, Seoul, South Korea.
Recruitment period	2010-2014.
Exclusion criteria	Participants with history of colonic examination, colorectal surgery, or CRN; inflammatory bowel diseases; ischemic or infectious colitis during the current colonoscopy; poor bowel preparation; and incomplete data for analysis.
Predictor measures	Age, sex, height, weight, family history of CRC, smoking habits, and comorbidities by using their electronic medical database.
Outcome measures	<i>ACRN was defined as cancer or advanced adenoma. Advanced adenoma was defined as the presence of one of the following features: diameter ≥ 10 mm, tubulovillous or villous structure, and high-grade dysplasia</i>
Statistical approach	Multivariable logistic regression.
Limitations	<ol style="list-style-type: none"> (1) The predicted probability cannot fully replace other screening modalities, including FIT, sigmoidoscopy, and colonoscopy. Although the predictive performance of our models was superior to that of existing models, it has been shown to be modest. (2) Computation of PAC-50 is complex; thus, it needs the use of a calculator. However, these calculators are available in most clinical settings. (3) Baseline characteristics significantly differ between the derivation and validation cohort. However, the prevalence of AN in both cohorts has meager differences. <p>Other limitations observed by the reviewer:</p>

- Use of split-sample technique may introduce potential bias in the model due to decreased sample size.
- Risk threshold determination was based solely on Youden index, which could decrease its potential clinical utility. Moreover, the model using a young population may not be very relevant in the context of cancer screening since the Asia-Pacific Screening recommendation did not yet approve the eligibility criteria of including participants > 50 years. Thus, the implementation of this model in a clinical setting, regardless of whether an external validation was/will be conducted, is not yet feasible.

New findings	In the young population, a predicted probability model can assess the risk of ACRN more accurately than existing models, including APCS, KCS, and Kaminski's scoring model.
---------------------	---

Jung 2018

Primary objective	To develop a risk-scoring model for predicting ACRN in FIT-negative persons.
Study location(s)	South Korea.
Recruitment period	2010-2014.
Population	FIT-negative persons.
Exclusion criteria	Individuals who have previous colonic examination, colorectal surgery, or colorectal neoplasia; a history of inflammatory bowel disease; ischemic or infectious colitis diagnosed during colonoscopy; poor bowel preparation; and incomplete data for analysis.
Predictor measures	Age; sex; height; weight; family history of CRC; smoking habits; drug history; and comorbidities, including hypertension, diabetes, dyslipidemia, old cerebrovascular attack, and fatty liver, were obtained from their electronic medical database.
Outcome measures	All participants were instructed to discontinue antiplatelet agents for 7 days and anticoagulants for 5 days, with the permission of the physician who prescribed the medication. The colonoscopic findings and results of the histopathologic examination were used to identify participants who had ACRN. Colonoscopy was performed by an experienced, board-certified endoscopist, using an Evis Lucera CV-260 colonoscope (Olympus Medical System, Tokyo, Japan). <i>CRN was defined as cancer or adenoma. ACRN was defined as cancer or advanced adenoma. Advanced adenoma was defined as the presence of one of the following features: diameter ≥ 10 mm, tubulovillous or villous structure, and high-grade dysplasia.</i>
Statistical approach	Multivariable logistic regression.
Limitations	(1) Selection Bias: Participants were sourced from just two medical examination centers in Korea, which may not represent the broader population. (2) Limited CRC Cases: The cohort included only 13 patients with CRC. As a result, the study focused on factors associated with ACRN rather than CRC. (3) Low-Risk Population: The study was based on a population with a low risk for CRC, thus limiting the model's generalizability to other demographic groups. (4) No Internal Validation: Due to the small size of the cohorts, the study could not conduct an internal validation of the model.
New findings	FIT-negative persons may need to undergo screening colonoscopy if they clinically have a high risk of ACRN. The scoring model based on age, smoking habits, overweight or obesity, hypertension, and old CVA may be useful in selecting and prioritizing FIT-negative persons for screening colonoscopy.

Kaminski 2014

Primary objective	To develop and validate a model to estimate the likelihood of detecting advanced colorectal neoplasia in Caucasian patients.
Study location(s)	Poland.
Recruitment period	January to December 2007.
Exclusion criteria	Those who were clinical suspicion of colorectal cancer; characteristics that met the criteria for Lynch syndrome, familial adenomatous polyposis, or inflammatory bowel disease; and colonoscopy within the preceding 10 years.
Predictor measures	Used self-administered questionnaire to collect information regarding the potential risk factors of ACN.

Outcome measures	Colonoscopy was conducted by an expert endoscopist. Colorectal finding was categorized based on the most advanced lesion identified at screening. <i>Advanced neoplasia was defined as cancer or adenoma that was at least 10 mm in diameter, had high-grade dysplasia, had villous or tubulovillous histological characteristics, or any combination thereof.</i>
Statistical approach	Multivariate logistic regression.
Limitations	The absence of a clear link between aspirin use and the risk of advanced colorectal neoplasia could be attributed to recall bias or incomplete data on the dosage and consistency of aspirin consumption.
New findings	The new score used age, sex, family history of colorectal cancer, cigarette smoking, and BMI to estimate the likelihood of detecting advanced colorectal neoplasia in asymptomatic Caucasian patients.

Kim 2019

Primary objective	To develop and validate a scoring system for ACRN in a large cohort comprising Korean subjects aged <50 years who underwent screening colonoscopy.
Study location(s)	South Korea.
Recruitment period	2003 to 2012.
Exclusion criteria	Incomplete colonoscopy, history of CRC or other cancers, history of IBD, history of previous colonoscopy, colorectal surgery, or missing clinical or laboratory data.
Predictor measures	Data on medical history, medication use, and health-related behaviors were collected using a self-administered questionnaire under the supervision of a well-trained interviewer.
Outcome measures	Colonoscopies were performed by experienced colonoscopists, who were unaware of the present study. Bowel preparations were performed using 4L of polyethylene glycol solution. Histological assessment of all polyps was performed by an expert pathologist who was unaware of the subjects' clinical data. <i>ACRN was defined as CRA ≥ 10mm in diameter, CRA with any component of villous histology, high-grade dysplasia, or carcinoma.</i>
Statistical approach	Use of univariable analysis to assess the potential predictors. Multivariable analysis with stepwise selection procedure based on Akaike information criterion for deriving the model.
Limitations	(1) The use of cross-sectional study design with a single ethnic group, which lowers the generalizability of the findings to other populations. Limitations identified by the reviewers: <ul style="list-style-type: none"> - The cross-sectional nature of the design could pose a potential bias due to lack of temporal information. The development of cancer is often a process that evolves over a period. Without longitudinal data, it is challenging to capture the progression of risk factors and their impact on cancer development. It could also lead to a misinterpretation of risk factors, leading to incorrect estimates of the model to predict the risk of CRC. - Survivorship bias: individuals might have already developed CRC and subsequently been treated or passed away, which could lead to biased estimates of survival probability. - Failure to account for this competing risk of "censoring bias."
New findings	The development of the Young Colorectal Screening (YCS) model used age, sex, alcohol consumption, smoking state, obesity, glucose metabolism abnormality, and family history of CRC in a large cohort of asymptomatic individuals aged < 50 years.

Liu 2018

Primary objective	To compare the performance of a simplified, largely categorized exposure-based colon cancer risk model against a more complex, largely continuous exposure-based risk model, using two prospective cohorts.
Study location(s)	United States.
Recruitment period	Enrollment in 1986.
Population	Female registered nurses aged 30 to 55 years and male health professionals aged 40 to 75 years. The analysis includes 63,219 women and 40,030 men, but the sample size is not estimated based assumption.
Exclusion criteria	Cancer diagnosis prior to 1986, had missing information on at least one of the colon cancer risk factors used in YDR, reported an unusual total energy intake (<500 kcal/day or >3500 kcal/day for women; <800

kcal/day or >4200 kcal/day for men), or had missing data of at least 10 items on the 1986 Food Frequency Questionnaire (FFQ).

Predictor measures	Height, body mass index (BMI), hormone replacement therapy (for women), physical activity, smoking, calcium intake from dairy food, alcohol and multivitamin intake, regular aspirin use, colonoscopy/sigmoidoscopy use, and family history of colorectal cancer in first-degree relatives.
Outcome measures	Incident cancer cases were identified through participants' self-reports on biennial follow-up questionnaires and confirmed through medical record review.
Statistical approach	Kaplan–Meier approach.
Limitations	<ol style="list-style-type: none"> (1) Self-Reported Risk Factors: The main limitation was the reliability of self-reported risk factors related to colon cancer, which can lead to inaccuracies. (2) Exclusion of Participants: Many participants with incomplete data on risk factors were left out of the analysis, potentially impacting the study's broad applicability. (3) Lack of Data on Aspirin Use: The study did not have detailed information on daily aspirin consumption in the HPFS, thus preventing the researchers from evaluating the long-term effects of aspirin on colon cancer risk and its influence on the model's efficacy. (4) Incomplete Endoscopic Screening Data: Details about the type of endoscopic colorectal cancer screening were missing for early follow-up periods up to 2002. This means that the defined variable for screening within a 10-year span might be underestimated, especially considering the recommended 5-year interval for sigmoidoscopy.
New findings	The results suggest that categorization of continuously distributed lifestyle and dietary factors did not significantly affect the discrimination and calibration of the model for colon cancer risk prediction.

Luu 2021

Primary objective	To validate APCS risk assessment tool for estimating the advanced colorectal neoplasia (ACN) risk at colonoscopy screenings and potential factors relevant for implementing this tool in the Korean population.
Study location(s)	National Cancer Center, South Korea.
Recruitment period	August 2002 to July 2014.
Population	People who visited the Center for Cancer Prevention and Detection at the National Cancer Center (NCC) for cancer screening, with 12,520 male and female were included for final analysis.
Exclusion criteria	Individuals younger than 40 years of age, individuals with a history of CRC, individuals for whom CRC was detected during a colonoscopy screening but who had another type of primary cancer, individuals with poor bowel preparation or who still had fecal matter in parts of the colon, and patients with incomplete information related to any variable.
Predictor measures	Age, sex, education, and household income, as well as family history of CRC, comorbidities, and health-related behavioral factors such as smoking status and drinking status. Body mass index (BMI) was calculated using weight and height, which were collected via a physical exam, using the formula.
Outcome measures	<p>The biopsy results of colonic polyps were provided in detail in the colonoscopy diagnosis, where the histological findings and the number and location of all polyps were noted by physicians based on the pathology report.</p> <p><i>Advanced colorectal neoplasia (ACN) cases were defined as either invasive cancer cases or advanced adenoma cases. Advanced adenomas were defined as polyps with a size of ≥ 10 mm, villous/tubulovillous histology, or high-grade dysplasia.</i></p>
Statistical approach	Logistic regression model.
Limitations	<ol style="list-style-type: none"> (1) Excluded Information: The study did not include certain colonoscopy-related factors (like adverse events, endoscopist experience, and history of colonoscopy) and dietary data in the analysis. (2) Quality of Procedures at NCC: Since the study was conducted at the NCC, Korea's premier cancer screening center, the standard of diagnosis, treatment, endoscopist proficiency, and colonoscopy quality may be higher than at smaller facilities. (3) Potential Self-Selection Bias: Participants were part of a cancer screening group and opted for colonoscopy-based CRC screening. This could indicate a bias, as these individuals might already belong to a higher socioeconomic class and be more health-conscious. (4) Single-Center Scope: The study's findings, derived from a single-center cohort, might not be wholly representative of the broader Korean population.
New findings	The APCS score could successfully classify Korean screens into different risk groups with acceptable discriminatory capability.

Ma 2010

Primary objective	To estimate a simplified score model used to estimate an individual's absolute CRC risk based on lifestyle information
Study location(s)	Japan Public Health Center-based (JPHC) Prospective study cohort II.
Recruitment period	1993 to 2005.
Population	a) Japanese patients; b) For development: 28,115 men. For validation: 18,256 men.
Exclusion criteria	Participants with a history of cancer or cardiovascular disease, diagnosed with cancers, or censored before the start of the follow-up survey.
Predictor measures	Use of self-administered questionnaire to obtain the demographic characteristics, including age and occupation, BMI, alcohol consumption, smoking status, diet, medical history, daily physical activity level, daily intake of nutrients, and other factors of the participants. Also, the Food Frequency Questionnaire, which is a validated tool, was used to measure the diet.
Outcome measures	Used population cancer registries. The cancer was coded using the International Classification of Diseases for Oncology for colon cancer and rectal cancer.
Statistical approach	Cox proportion hazard regression.
Limitations	(1) The score model included components based on calculations, e.g., alcohol consumption (gram/week) and physical activity (MET-hour/day). (2) Although the model was externally validated by an independent cohort, risk factor profiles and measurement were like those of the population for model development. Other limitations or potential bias observed the reviewer: <ul style="list-style-type: none"> - Use of univariate analysis to identify potential predictors might lead to overestimation or underestimation of the importance of certain predictors, as it does not account for confounding variables. - Exclusion of highly relevant population due to its comorbidities might bias the estimates of the model. Although the researchers used Cox regression to estimate the 10-year absolute risk of developing CRC, they did not account for competing risk and did not account for the "censoring" bias.
New findings	The model identified age, alcohol consumption, and daily activity level to be the most important CRC risk factors.

Meester 2022

Primary objective	To evaluate the use of F-Hb in prediction models.
Study location(s)	Dutch biennial FIT-based screening program, the Netherlands.
Recruitment period	2014-2019.
Population and sample size	Among 265 881 participants that completed three rounds of FIT, 8806 had a positive FIT result.
Exclusion criteria	Participants with a positive FIT in the third round without complete follow-up colonoscopy, no age-eligible individuals.
Predictor measures	Age, sex, and F-Hb (first and second times).
Outcome measures	Outcome Assessment: In the Dutch CRC screening program, adults between 55 and 75 years are given an FIT test via mail. Positive FIT results lead to a colonoscopy. During the colonoscopy, polyps are removed, and necessary biopsies are conducted. Positive FIT individuals typically receive a colonoscopy within 15 days. Quality is upheld by certified endoscopists, and if a colonoscopy is unclear due to bowel preparation issues, a re-examination is performed. Detected lesions are reviewed by pathologists. All labs and centers undergo annual quality audits. <i>Relevant outcomes in the program are advanced adenomas (criteria: size ≥ 10 mm, villous histology $\geq 25\%$, or high-grade dysplasia presence) and colorectal cancer (CRC). Those identified with these findings may need further treatment or monitoring as per Dutch guidelines, while those without return for FIT screening in 10 years.</i>
Statistical approach	Multivariate logistic regression.
Limitations	(1) Representativeness: Excluding participants with high CRC risk and variations in FIT cutoff values across countries may affect the generalizability of the study.

- (2) Incomplete Data: The study considered only two rounds of F-Hb, but a sensitivity analysis indicated more CRC diagnoses after a third round.

New findings	The findings indicated that these models not only accurately predicted the risk of subsequent AN and CRC but also effectively discriminated between these outcomes. Based on these promising results, the study suggests that clinical practice should consider adopting risk-stratified FIT screening methodologies.
---------------------	---

Müdler 2023

Primary objective	We validated this existing logistic regression (LR) model and attempted to improve it by applying a more flexible machine-learning approach.
Study location(s)	The Netherlands.
Recruitment period	2014 to 2021.
Population and sample size	All Dutch citizens aged between 55 and 75 years. Sample of 1,356,860.
Exclusion criteria	We excluded participants with missing FITs in round 1 and 2, positive FITs in round 1 and 2, missing findings in the participant records on sex and/or age and participants with a positive FIT but no follow up colonoscopy in round 3.
Predictor measures	Age, sex, and the two most recent f-Hb concentrations.
Outcome measures	During the follow-up colonoscopy, advanced neoplasia (AN) is considered to be a relevant finding. AN consists of the presence of either advanced adenomas or CRC.
Statistical approach	Logistic regression (LR) model and random forest (RF) model.
Limitations	<ol style="list-style-type: none"> (1) Validation Scope: The study's validation was carried out within the same screening program where the LR model was developed, limiting external generalizability. (2) Lack of Interval Cancer Data: Data regarding interval cancers were not available during the study. Including such data would enhance the study's validity. (3) Validation Rounds: The study's validation was constrained to the third and fourth rounds of the screening program. Validating the study with future rounds would affirm its legitimacy and show the value of accumulating data over time. (4) Temporal Nature and Imbalance: The models presented did not adequately address the data's temporal aspects and imbalances.
New findings	An RF model does not improve CRC risk prediction compared to an LR model, probably due to the limited number of available explanatory variables. Therefore, the LR remains the preferred prediction tool because of its interpretability.

Murchie 2017

Primary objective	To develop a prediction model for high-risk colon adenomas in an average-risk population.
Study location(s)	Cleveland Clinic Florida, United States.
Recruitment period	August 2008 to August 2014.
Population	5063 patients aged from 40 to 59 years.
Exclusion criteria	Patients with a first-degree family history of CRC, personal history of inflammatory bowel disease, previous colonic resection, weight loss, gastrointestinal bleeding, iron deficiency anemia as indications for colonoscopy, and incomplete colonoscopies.
Predictor measures	Age, body mass index, sex, race, and smoking history.
Outcome measures	Colonoscopies were performed by gastroenterologist or fellow under attending supervision. Endoscopy report included depth of insertion and quality of preparation as required fields for report completion. A complete colonoscopy was defined as cecal intubation (recognition of the appendiceal orifice, ileo-cecal with associated photograph). (High-risk colon adenomas: ≥ 3 nonadvanced adenomas.)
Statistical approach	Penalized logistic regression using bootstrap validation.
Limitations	<ol style="list-style-type: none"> (1) Validation Concerns: The study lacked external validation, which could raise questions about its broader applicability.

- (2) Smoking History: The study could not quantify the smoking history, offering only binary options (yes/no), without detailing the extent or duration of smoking.
- (3) Excluded Variables: Due to its retrospective design, the study omitted certain factors like alcohol consumption, dietary habits, and physical activity.
- (4) Age Factor: The participants' ages were not aligned with typical indications for colonoscopy screenings.

New findings	This model was compared with 4 models with external validation, namely the Kaminsky model, Yeoh model, Driver model, and Betés models, and was found to be more favorable than all of them. Age and BMI predictors were used as continuous variables. It also reported absolute predictive probabilities of advanced and high-risk polyps, allowing for a more individualized risk assessment of CRC.
---------------------	---

Musselwhite 2019

Primary objective	To externally validate the National Cancer Institute CRC risk Assessment tool, which calculates the future risk of CRC, to see if it could be used to predict current AN in veteran cohort undergoing baseline screening colonoscopy.
Study location(s)	13 diverse VA medical centers in United States.
Recruitment period	1994 to 1997.
Population	Veterans aged 50 to 75 years.
Exclusion criteria	Those who have gastrointestinal disease, lower endoscopy in previous 10 years, colon surgery, significant co-morbidity, or other medical condition that could increase the risk of performing a screening colonoscopy.
Predictor measures	Used validated, detailed questionnaire to obtain medical history and lifestyle collected 6 months before screening colonoscopy.
Outcome measures	Centrally trained pathologists blinded participants to information reviewed biopsies at the site of care. Biopsies were then sent for a blinded second review, with discrepancies resolved by a third referee pathologist. <i>ACN was defined as the presence of ≥ 1 cm, villous histology, high-grade dysplasia, or carcinoma.</i>
Statistical approach	Wilcoxon rank-sum test to test the null hypothesis of no difference in median risk scores among AN cases and non-cases. A logistic regression model was used to evaluate the model performance, using AUC.
Limitations	<ol style="list-style-type: none"> (1) Demographic Limitation: The study mainly involved male veterans from the 1990s, making it hard to evaluate its applicability to women. (2) Missing Data: Waist circumference was not measured, thus hindering comparison to other relevant models. (3) Repeat Screenings: The tool's effectiveness in repeated screenings or follow-ups is undetermined. (4) Data Accuracy: There may be inaccuracies in the data regarding diet, medication, and other personal factors. (5) Score Overlap: Scores for those with and without advanced conditions overlapped, making risk differentiation challenging. (6) Tool's Narrow Scope: This tool should not be the only basis for CRC screening decisions. Other factors, like patient preference and healthcare capacity, also matter. (7) Future Enhancements: Uncertain if adding genetic data will improve the tool's effectiveness.
New findings	The tool has a modest discriminatory function for estimating the presence of current advanced neoplasia in veterans undergoing a first screening colonoscopy. These findings are comparable to other clinically utilized cancer risk prediction models and may be used to inform the benefit-risk assessment of screening, particularly for patients with competing comorbidities and lower risk, for whom a noninvasive screening approach is preferred.

Park 2019

Primary objective	To develop risk stratification models for ACN and CRC based on fecal hemoglobin (f-HB) concentration and clinical risk factors.
Study location(s)	National Cancer Screening Program, Korea.
Recruitment period	May 2013 to April 2017.

Population	Retrospectively included asymptomatic participants who underwent FIT and colonoscopy for CRC screening.
Exclusion criteria	Aged <50 years, family history of CRC, previous history of CRC or colorectal surgery, history of inflammatory bowel diseases, and poor bowel preparation.
Predictor measures	Clinical risk factors were obtained, including age, sex, family history of CRC, smoking habits, and body mass index, from an electronic medical database. Data on the family history of CRC and smoking habits were collected using a self-administered questionnaire.
Outcome measures	<p>Stool samples were obtained using a sampling tube (Eiken Chemical Company, Tokyo, Japan) containing 2.0 mL of buffer designed to minimize hemoglobin degradation. The collected fecal material was sealed in a plastic bag and sent to the laboratory.</p> <p>The f-Hb quantitation was performed using OC-Sensor Diana.</p> <p>All colonoscopies were performed by experienced board-certified endoscopists, using an Evis Lucera CV-260.</p> <p><i>Advanced adenoma was defined as the presence of 1 of the following features: tumor diameter ≥ 10 mm, tubulovillous or villous structure, and high-grade dysplasia.</i></p>
Statistical approach	Multivariable logistic regression.
Limitations	<ol style="list-style-type: none"> (1) Only one type of FIT device, the OC-Sensor, was used, and the FIT cutoff used could be different from other clinical settings. (2) Lack of external validation. (3) Exclusion of individuals with family history of CRC, which limits the model's generalizability to this group. (4) This model did not consider potential confounding factors including use of nonsteroidal anti-inflammatory drugs, alcohol history, and abdominal radiation exposure history because information on these factors was not available in their retrospective cohort. <p>Other limitations identified by the reviewer:</p> <ul style="list-style-type: none"> - Use of univariate analysis to identify potential predictors might lead to overestimation or underestimation of the importance of certain predictors, as it does not account for confounding variables. - Did not account for competing risk, timing of outcome and predictor assessment, or blinding of assessors for both predictor variables and outcomes, thus possibly leading to biased estimates of the model. - Unclear justification for risk threshold determination and lack of reported estimates on the potential clinical utility of the model. - Small sample size due to significant number of participants excluded.
New findings	The proposed model can effectively stratify the risk for ACRN and CRC and provide accurate information on this risk in individuals who undergo FIT.

Ruco 2015

Primary objective	External validation of a previous risk index for advanced neoplasia.
Study location(s)	Women's College Hospital in Toronto and Alberta Health Service's Colon Cancer Screening Centre, United States.
Recruitment period	2003 to 2008.
Population	5137 asymptomatic participants with mean aged 58.3 years.
Exclusion criteria	<ol style="list-style-type: none"> (1) Previous history of colon surgery. (2) History of ulcerative colitis, colon polyps, or colon cancer. (3) Rectal bleeding in the previous six months on more than one occasion. (4) Marked change in bowel habits in the previous six months. (5) Lower abdominal pain that would normally require medical attention in the previous six months. (6) Previous history of sigmoidoscopy, colonoscopy, or barium enema within the past 10 years. (7) A disease that would preclude the safe performance of colonoscopy.
Predictor measures	Kaminsky score: age, sex, family history, smoking history in pack-years, and BMI.
Outcome measures	The outcome was measured using a colonoscopy. Findings were categorized based on the most advanced finding. <i>Advanced neoplasia (AN) included cancer or a tubular adenoma, traditional serrated adenoma (TSA), or sessile serrated adenoma (SSA) with villous characteristics ($\geq 25\%$ villous component), and/or high-grade dysplasia and/or diameter ≥ 10 mm.</i>

Statistical approach	Multivariate logistic regression.
Limitations	(1) Questionnaire Adaptation: The questionnaire, originally tailored for the Canadian population, may not have perfectly aligned questions. (2) Evaluation Concerns: The study might have a reduced capacity to assess the index's performance for high-risk individuals due to fewer cases.
New findings	This predictive model was less predictive of advanced neoplasia in comparison with the original study. It confirmed the association between smoking, BMI, and AN.
Additional notes:	The study is based on the Kaminsky predictive model.

Schroy III 2015

Primary objective	To develop and validate a clinical index for estimating the probability of advanced colorectal neoplasia at screening colonoscopy.
Study location(s)	Boston Medical Center and Tufts Medical Center, New England.
Recruitment period	March 22, 2005, to January 31, 2012.
Population	Asymptomatic, mostly English-speaking participants, aged 50-79 years.
Exclusion criteria	Patients with incomplete examinations due to poor bowel preparation or failure to reach the cecum for reasons other than a poor bowel preparation obstructing neoplasms were excluded from the analysis if they did not undergo a complete examination.
Predictor measures	The risk assessment questionnaire (21-item) was self-administered to consenting patients with adequate literacy skills, using a scannable, paper-based data-collection form. A trained interviewer technique was used for patients with low literacy skills.
Outcome measures	All colonoscopies were performed by board-certified attending gastroenterologists alone or assisted by a gastroenterology fellow. All retrieved polypoid lesions or biopsy specimens were reviewed initially by board-certified pathologists and classified according to the World Health Organization's histological criteria. <i>An advanced colorectal neoplasm was defined as a tubular adenoma ≥ 10 mm in size, an adenoma of any size with villous features or high-grade dysplasia, a dysplastic serrated lesion of any size, or invasive cancer.</i>
Statistical approach	An expectation-maximization algorithm was used to obtain estimates of the variance-covariance matrix and model coefficients for logistic regression. Used multiple logistic regression, using backward selection on imputed dataset.
Limitations	(1) Lack of External Data: No external dataset was used. (2) Selection Bias: Patients were consecutively recruited, with a high enrollment rate and a reliance on a convenience sample. (3) Exclusion of Certain Parameters: The study did not account for measures like hip-to-waist circumference or C-reactive protein levels. (4) Potential Misclassification: Subjective judgments by multiple endoscopists regarding polyp size could lead to misclassification of ACN based solely on size. (5) Social-Response Bias: Self-reported data, especially concerning BMI, alcohol consumption, and smoking, might not be accurate. (6) Model's Scope: The model was built on data from patients willing to have a screening colonoscopy. Its effectiveness for those unwilling or unable to undergo such a procedure remains untested.
New findings	The final index consisted of 5 independent predictors, namely age, smoking, alcohol intake, height, and a combined sex/race/ethnicity. The performance varied based on sex and race/ethnicity but could accurately stratify average-risk patients into low- and intermediate/high-risk categories for CRC at the screening colonoscopy.

Sekiguchi 2018

Primary objective	To build a new useful scoring model for CRC screening, externally validate the modified APCS score, and compare the usefulness of two scores.
Study location(s)	National Cancer Center, Tokyo, Japan.
Recruitment period	February 2004 to March 2013.

Population	Asymptomatic Japanese individuals who underwent first screening colonoscopy.
Exclusion criteria	Unavailability of data from self-administered questionnaire on lifestyle, demographic characteristics, and medical history, which all screened individuals, poor bowel preparation, and refusal to participate in the study.
Predictor measures	Risk factors were retrospectively assessed using the data from all included study participants.
Outcome measures	Colonoscopy was used to identify ACN by experienced endoscopists certified by the Japanese Gastrointestinal Endoscopy Society. <i>Advanced neoplasia was defined as a tubular adenoma or serrated lesion ≥ 10 mm in size, any adenoma with villous features, or any lesion with high-grade dysplasia or carcinoma.</i>
Statistical approach	Multivariate logistic regression with bootstrapping for internal validation of the derived model.
Limitations	<ol style="list-style-type: none"> (1) Data of ACN and adenomatous lesions were based on the results of a single examination; thus, some lesions may have been overlooked. (2) Although the newly developed scoring model was internally validated, external validation was lacking. <p>Other limitations identified by the reviewers:</p> <ol style="list-style-type: none"> (3) Use of univariate analysis to identify potential predictors could lead to potential bias in identifying relevant predictors due to confounding factors. (4) Lack of reported estimates on the observed/expected ratio of the model, as well as other parameters (e.g., sensitivity, specificity, negative predictive value and/or positivity predictive value, and net benefits), thus limiting the clinical applicability, interpretability, and transferability of the model to clinical settings.
New findings	An 8-point scoring model for the prediction of ACM was developed using five independent risk factors.

Sharara 2020

Primary objective	To create and internally validate a risk prediction model for the detection of advanced neoplasia in average-risk individuals.
Study location(s)	American University of Beirut Medical Center, United States.
Recruitment period	5-year period (no mention of year collected).
Population	Average-risk asymptomatic patients who were scheduled for screening colonoscopy.
Exclusion criteria	A prior history of colonoscopy, known colon polyps, inflammatory bowel disease, had undergone previous colonic resection, had a family history of CRC, and AN in any first-degree relative or two or more second-degree relatives at any age. Diagnostic colonoscopies performed for symptoms such as bleeding or abdominal pain.
Predictor measures	Conducted interviews using a paper-based questionnaire to obtain BMI, smoking (pack years), age, and daily red meat consumption prior to the procedure (colonoscopy). The questionnaire included 18 factors, including demographics; tobacco and alcohol use; dietary patterns; and concomitant medications and supplements such as aspirin, NSAIDs, oral contraceptive pills/hormone replacement therapy, and calcium supplements.
Outcome measures	Information on withdrawal time; quality of bowel preparation; and location, size, number, and histology of polyps was collected. <i>Advanced neoplasia was defined as a tubular adenoma or serrated lesion ≥ 10 mm in size, any adenoma with villous features, or any lesion with high-grade dysplasia or carcinoma. In cases of multiple polyps, classification was based on the most advanced histology.</i>
Statistical approach	Multivariate binary logistic regression using backward stepwise approach and bootstrapping for internal validation of the model.
Limitations	<ol style="list-style-type: none"> (1) External validation in a separate population is needed. (2) The study included patients with private insurance. Thus, patients are not entirely represented. (3) A probable underestimation of the risk effect due to the population was derived from patients willing to undergo a screening colonoscopy. <p>Other limitations identified by the reviewers:</p> <ol style="list-style-type: none"> (2) Use of univariate analysis to identify potential predictors could lead to potential bias in identifying relevant predictors due to confounding factors. (3) Lack of reported estimates on the observed/expected ratio of the model, as well as other parameters (e.g., sensitivity, specificity, negative predictive value and/or positivity predictive

value, and net benefits), thus limiting the clinical applicability, interpretability, and transferability of the model to clinical settings.

New findings	Age as a risk factor for AN shows an additive effect when combined with other risk factors such as high BMI and smoking.
---------------------	--

Shin 2014

Primary objective	To develop colorectal cancer risk prediction models for overall colorectal proximal colon, distal colon, and rectal cancer.
Study location(s)	National Health Insurance Corporation, South Korea.
Recruitment period	1996 to 1997 development participants. 1998 to 1999 validation participants.
(a) Population (b) Sample size	(a) Korean participants; (b) 1,326,058 (846,559 men and 479,499 women) for development set, and 963,749 (547,874 men and 415,875 women) for validation set.
Exclusion criteria	Cancers with an overlapping lesion of the colon (C188), and those that were not otherwise specified (C189) were excluded from the analysis.
Predictor measures	Self-administered questionnaires were used to collect risk factors, including age, serum cholesterol, alcohol consumption, weight, height, cigarette smoking habits, regular exercise, family history of CRC, dietary preferences, information about female reproductive factors, and meat intake frequency.
Outcome measures	Cancer was ascertained from the Korean Central Cancer Registry (KCCR) database, and death information from the Korean National Statistical Office. The subsites of colorectal cancer were categorized by the International Classification of Disease 10th edition (ICD-10) code as follows: <i>proximal colon (C180–C185), distal colon (C186–C187), and rectum (C19–C20)</i> .
Statistical approach	Cox proportional-hazard regression for developing prediction equations. Five models were developed for overall colorectal cancer, colon cancer, right colon cancer, left colon cancer, and rectal cancer, separately for men and women.
Limitations	(1) Limited information on dietary risk or protective factors such as calcium and fiber intake, or non-dietary factors such as nonsteroidal anti-inflammatory drugs. Other limitations identified by the reviewers: (2) Use of univariate analysis to identify potential predictors could lead to potential bias in identifying relevant predictors due to confounding factors. (3) Lack of reported estimates on sensitivity, specificity, negative predictive value and/or positivity predictive value, and net benefits, thus limiting the clinical applicability, interpretability, and transferability of the model to clinical settings. (4) Absolute risk probability was not reported, and neither was risk stratification, which could also raise concern on its potential clinical utility.
New findings	Age, BMI, serum cholesterol, family history of CRC, and alcohol consumption were included in all models for men, while age, height, and meat intake frequency were included in all models for women.

Soonklang 2021

Primary objective	To compare model performance of binary logistic regression (BLR), polytomous logistic regression (PLR), and classification and regression tree (CART) between the clinical prediction scores of advanced colorectal neoplasia in asymptomatic Thai patients.
Study location(s)	Royal Charity Project, Thammasat University, Thailand.
Recruitment period	July 2009 and June 2010.
Population	Asymptomatic Thai patients.
Exclusion criteria	Patients with a history of colorectal cancer or colonoscopy within 10 years were excluded from the study.
Predictor measures	Use of standard questionnaire to obtain risk factors, including sex, age, BMI, family history of CRC in first-degree relatives, alcohol use, smoking history, and diabetes mellitus. Unclear assessment of fecal immunochemical test (FIT) results, as they were not reported in the article.
Outcome measures	The participants underwent colonoscopy screening. Patients were categorized into 2 groups based on colonoscopic findings and pathological reports. <i>The first group was ACN with malignant, villous, or tubulovillous histologic characteristics, high-grade dysplasia, or adenomatous lesions ≥ 10mm in diameter. The second group included other polyps with pathological reports: adenoma size <1 cm, hyperplastic polyps, inflammatory polyps, colitis, lipoma, and no colorectal tumor.</i>

Next, the patients were categorized into 3 groups based on colonoscopy findings and pathological reports. The first group comprised those with ACN *with malignant, villous, or tubulovillous histologic characteristics; high-grade dysplasia; or adenomatous lesions ≥ 10 mm in diameter. The second group (moderate group) comprised those who had other polyps with pathological reports of adenoma size < 1 cm and hyperplastic polyps ≥ 1 cm.*

The third group (average group) comprised *those who had hyperplastic polyps < 1 cm in size, inflammatory polyps, colitis, lipoma, and no colorectal tumors.*

Statistical approach	BLR, PLR, and CART with bootstrapping for internal validation of each model. Comparison of predictive scores for ACN was performed using AUROC and chi-square to compare AUROC scores.
Limitations	(1) The data were only secondary data from the Royal Charity Project of Colorectal Cancer Screening. (2) Additional details of clinical information or clinical risks, such as waist circumference, may increase the accuracy of this risk score. (3) The validation of the study was not prospectively collected. Other limitations identified by the reviewers: (4) Use of univariate analysis to identify potential predictors could lead to potential bias in identifying relevant predictors due to confounding factors. (5) Lack of reported estimates on sensitivity, specificity, negative predictive value and/or positivity predictive value, and net benefits, thus limiting the clinical applicability, interpretability, and transferability of the model to clinical settings. (6) The estimates for the E/O ratio were unclear, and neither the absolute risk probability nor the risk stratification was reported, which could also raise concern on its potential clinical utility.
New findings	The BLR and CART models yielded similar accuracies for the prediction of ACN in Thai patients. The PLR model provided higher accuracy for ACN prediction than the CART model.

Stegeman 2014

Primary objective	To explore the use of a risk prediction model in CRC screening.
Study location(s)	Regional Comprehensive Cancer Center in Amsterdam and Rotterdam, the Netherlands.
Recruitment period	2013.
Population	Asymptomatic men and women between 50 years and 75 years of age, 6600 asymptomatic men and women. Validation not performed.
Exclusion criteria	Subjects who were in a colonoscopy surveillance program and those with a life expectancy less than 5 years.
Predictor measures	A self-administered questionnaire was handed out to participants in the waiting room before colonoscopy. Predictors collected include age, family history (first degree) of CRC, alcohol intake, current smoking, history of smoking, BMI, regular aspirin or non-steroid anti-inflammatory drug (NSAID) use, total calcium intake, and physical activity. Regular NSAID intake was defined as the use of NSAIDs three or more times a week during the last month. Calcium intake was estimated by questions about food and supplement intake.
Outcome measures	Colonoscopy was performed using the standard quality aspects defined by the American Society for Gastrointestinal Endoscopy. Most advanced lesion per patient was used. Advanced neoplasia is the outcome. <i>Advanced neoplasia was defined as at least one CRC or advanced adenoma: adenoma of 10 mm or larger, $\geq 25\%$ villous histology or high-grade dysplasia.</i>
Statistical approach	Multivariate logistic regression model using backward elimination to develop a parsimonious model with a significance level of 0.2 as the removal criterion. Penalized shrinkage was used to correct for optimism.
Limitations	(1) Low response rate for colonoscopy; participation rates in primary colonoscopy screening are generally suboptimal.
New findings	Adding risk-based stratification increases the accuracy FIT-based CRC screening and could be used in preselection for colonoscopy in CRC screening programs.

Sung 2018

Primary objective	To validate a modified risk algorithm based on the Asia-Pacific Colorectal Screening (APCS) score that included body mass index (BMI) for prediction of advanced neoplasia.
--------------------------	---

Study location(s)	Hong Kong.
Recruitment period	2008 to 2012.
Exclusion criteria	Individuals who have personal history of CRC, colonic adenoma, inflammatory bowel disease, prosthetic heart valve, or vascular graft surgery, as well as medical conditions, which were contraindications for colonoscopy.
Predictor measures	Use of self-administered questionnaire, consisting of details on their sociodemographic and clinical information. The bodyweight and height were measured with the participant wearing light clothing, without wearing shoes, using wall mounted stadiometer and regularly calibrated weight scales.
Outcome measures	Use of a standardized bowel preparation called polyethylene glycol (Klean-Prep [®] , Helsinn Birex Pharmaceuticals Ltd. Ireland). Colonoscopy was performed by experienced colonoscopists in an endoscopy center. Definition of the outcome—advanced neoplasia not reported.
Statistical approach	Binary logistic regression.
Limitations/Bias	<ol style="list-style-type: none"> (1) Selection Bias and Misclassification: Excluding individuals with certain medical conditions or contraindications might result in an unrepresentative sample and potential misclassification of outcomes. (2) Ethical Concerns: The exclusion of individuals based on medical conditions, particularly if these conditions are common among disadvantaged or vulnerable groups, raises ethical issues. (3) Risk Underestimation: The true risk might be underestimated due to the exclusion of certain individuals.
New findings	Incorporating BMI into the predictors of APCS score was found to improve risk prediction of advanced neoplasia and reduce colonoscopy resources.

Sutherland 2021

Primary objective	To develop a risk prediction model for high-risk AN detected at screening colonoscopy based on readily available participant information.
Study location(s)	Endoscopy unit in Calgary, AB, Canada.
Recruitment period	2008 to 2016.
Population	Participants aged 50 to 74 years with no history of cancer, free of any major comorbidities and significant colorectal symptoms that would preclude non-hospital-based endoscopy. Moreover, only those who underwent first time screening colonoscopy were included.
Exclusion criteria	Those with personal history of CRC, polyps, or adenomas, or those showing irritable bowel disease symptoms were excluded.
Predictor measures	Exposure variables were collected prior to colonoscopy via a health and lifestyle questionnaire that collects both sociodemographic and clinical information.
Outcome measures	Primary outcome of interest was the presence of high-risk AN, defined as <i>an adenoma with villous histology, high-grade dysplasia, ≥10 mm in diameter, or 3 or more adenomas</i> . This criterion was ascertained by endoscopist report form completed by the physician following the procedure and subsequent pathology report.
Statistical approach	<ol style="list-style-type: none"> (1) Variables included in the model were selected based on a priori according to epidemiological evidence regarding the factors associated with HRAs and CRC. (2) Multiple logistic regression, accounting for the possibility of multicollinearity, using ridge regularization, to estimate the coefficients. (3) Bootstrapping was used to internally validate the derived model. (4) Missing data were handled using stochastic regression imputation.
Limitations	<ol style="list-style-type: none"> (1) Although only seven predictors were used in the model, there is still a requirement of a clinician to ascertain and have low confidence in the information being included in the model. (2) Several predictor variables used, such family history of CRC and alcohol consumption, are subject to reporting and misclassification biases, as well as recall (for family history) and social desirability bias (alcohol consumption). (3) Potential population referred to the center makes healthier lifestyle choices than those of the general population. (4) Lack of external validation to an independent cohort; thus, it would be unlikely to contribute to CRC screening meaningfully in its current state.
New findings	The internally validated risk prediction model yielded an optimistic AUC of 67% and displayed strong specificity and NPV for the detection of high-risk adenoma.

Tao 2014

Primary objective	To develop and validate a scoring system to identify individuals at high risk for advanced colorectal neoplasms among the average-risk population for CRC screening based on some easy-to-collect risk factors.
Study location(s)	Germany.
Recruitment period	Derivation sample 2005-2009. Validation sample 2005-2011.
(a) Population (b) Sample size	(a) German screening program; (b) Derivation set: 7891. Validation set: 3519.
Exclusion criteria	Participants with a history of CRC or previous colorectal surgery, participants who reported having undergone a previous colonoscopy within the past five years, and those with poor bowel preparation in the endoscopy report were excluded in the analysis.
Predictor measures	Used a standardized questionnaire to obtain information on sex, age, first-degree relatives with a history of CRC, cigarette smoking, alcohol consumption, red meat consumption, ever regular use (at least 2 times/week for at least 1 year) of nonsteroidal anti-inflammatory drugs, previous colonoscopy, and previous detection of polyps.
Outcome measures	Colonoscopy conducted by an experienced endoscopist. <i>Advanced colorectal neoplasms include CRC and advanced colorectal adenomas, which are defined as adenomas with at least one of the following features: 1 cm or larger in size, tubulovillous or villous components, or high-grade dysplasia</i>
Statistical approach	Multivariate logistic regression model.
Limitations	<ol style="list-style-type: none"> (1) Inaccuracies in self-reported risk factor information, such as for smoking or alcohol consumption. (2) Improper detection of neoplasms at colonoscopies. (3) A majority of patients were Caucasian European, thus making it difficult to generalize. (4) Potential selection bias due to the participation rate in the CRC screening program is still relatively low in Germany. <p>Other limitations observed by the reviewers:</p> <ol style="list-style-type: none"> (5) The split-sampling technique may be an inefficient approach to internal validation. (6) Transformation of continuous predictors to categorical variables may raise concern about the loss of valuable information of the data. (7) Risk threshold determination was arbitrarily defined, which may not be applicable in clinical context.
New findings	A scoring system was developed based on two large cohorts showing promising performance. The model comprised 9 risk factors that were significantly associated with AN.

Thomsen 2022

Primary objective	To develop and validate a risk stratification model to calculate predicted risk based on information available at the time of FIT participation (age, gender, and fecal hemoglobin (fHb) values). To compare how many CRCs and adenomas each method (risk vs. FIT cutoff) would identify or miss with reduced number of colonoscopies.
Study location(s)	Denmark.
Recruitment period	2014 to 2016.
Population	Cohorts from the Danish Colorectal Cancer Screening Database (DCCSD) and the Danish Colorectal Cancer Group Database (DCCG). Restricted participants to persons who underwent colonoscopy within 90 days of positive test for FIT results.
Exclusion criteria	Not reported.
Predictor measures	Age was retrieved at the time of a positive test, gender, and fHb value from the DCCSD. Age was categorized in 5-year intervals. Fecal hemoglobin level was also categorized to 20-29, 30-44, 45-79, 80-199, and $200 \leq \mu\text{g fHb/g feces}$.
Outcome measures	Findings at colonoscopy were categorized as no findings; low-risk, medium-risk, or high-risk adenomas; or CRC. <i>CRC was defined as a diagnosis recorded in either the DCCSD or in the DCCG within 90 days of a positive FIT. In the DCCG, low-risk adenomas were defined as less than three adenomas less than 10mm or a tubular adenoma or at least one adenoma with low-grade dysplasia. Medium-risk adenomas were defined as three or four adenomas,</i>

or at least 10-19 mm in size, or least one tubulovillous or villous adenoma, or at least one adenoma with high-grade dysplasia. High-risk adenomas were defined as five or more adenomas, or at least one adenoma ≥ 20 mm in size, or at least one adenoma removed by piecemeal technique.

Statistical approach	Multiple regression analysis with bootstrapping and temporal validation, using 2016 dataset.
Limitations	Not reported Limitations identified by the reviewers: <ol style="list-style-type: none"> (1) The inclusion criteria strictly involved individuals who tested positive on the FIT, which limits the generalizability of the model to a broader population, particularly those who tested negative on the FIT or those who chose not to undergo FIT testing altogether. Potential high-risk populations could still be present within the FIT-negative population due to the limited predictive accuracy of the test. As a result, the study might have excluded eligible populations for colonoscopy based on a FIT-negative result or the decision not to undergo FIT testing at all. (2) Use of temporal validation poses a risk of over-optimistic results due to data similarity in the derivation and validation sets and may affect the generalizability of the model to different sets of cohorts.
New findings	Their prediction model permits predicted risk to determine colonoscopy referral, instead of a dichotomized FIT result alone.

Van 't Klooster 2020

Primary objective	To derive and externally validate prediction models for the estimation of lifetime and 10-year risk for total, colorectal, and lung cancer in patients with established cardiovascular disease.
Study location(s)	Utrecht, the Netherlands.
Recruitment period	Not reported.
Population	Cohorts from the UCC-SMART (Utrecht Cardiovascular Cohort-Second Manifestations of ARterial disease) study were used for model development and the CANTOS (Canakinumab Anti-Inflammatory Thrombosis Outcomes Study) trial for the model validation.
Exclusion criteria	None.
Predictor measures	Age, sex, smoking status, weight, height, alcohol use, use of antiplatelet medication, and diabetes mellitus were selected based on the readily clinically available potential predictors that were present in both the derivation and validation sets. A biannual questionnaire was provided to gather information about the occurrence and recurrent of CVD, bleeding events, incident diabetes mellitus, and end-stage renal disease.
Outcome measures	For colorectal cancer outcome, the 10-year and absolute lifetime risk of developing the targeted cancer was estimated. Cancer outcomes were obtained from the national registry receiving notifications of all new cancer diagnoses, especially lung and colorectal cancer, the Registry of Histopathology and Cytopathology, and hospital discharges diagnoses. Competing risk: (1) noncancer death, (2) non-colorectal-cancer death, and (3) non-lung-cancer death.
Statistical approach	Fine and Gray competing risk-adjusted sub-distribution hazard functions developed with left truncation: age rather than follow-up time was used as the underlying time scale. Coefficient of predictors were adjusted to account for optimism using a shrinkage factor acquired by bootstrapping samples.
Limitations	<ol style="list-style-type: none"> (1) Smaller number of lung and colorectal cancer in the development and validation sets. (2) External validation in the CANTOS trial could only be performed up to 4 years due to limited follow-up duration, although UCC-SMART's internal validation for the 10-year prediction showed good calibration. (3) Several potentially important predictors, including level of education, socioeconomic status, race, and family history of cancer, were unavailable in the derivation cohort and could not be included in the prediction model, possibly limiting the model performance.
New findings	The lifetime and 10-year risk of total cancer, colorectal cancer, and lung cancer can be estimated reasonably well with easily clinically available predictors in patients with established CVD. The wide distribution of predicted lifetime risks for total and lung cancer enables identification of patients at the highest risk for cancer.

Yang 2017

Primary objective	To derive and validate a novel risk-scoring model to predict the risk of ACN in a Korean population.
Study location(s)	Kangbuk Samsung Hospital Health Screening Center, Seoul, South Korea.
Recruitment period	January 2003 to December 2012.
Population	Asymptomatic adult recipients of a screening colonoscopy during a comprehensive health checkup program.
Exclusion criteria	Those who have previous colorectal examination, history of CRC or other malignancy, history of inflammatory bowel disease, history of colorectal surgery, incomplete colonoscopy or failed bowel preparation, or those with missing clinical or laboratory data.
Predictor measures	Medical history and lifestyle factors were determined via a self-administer questionnaire.
Outcome measures	Colonoscopies were performed by 13 experienced colonoscopists, with Evis Lucera CV-260 colonoscope (Olympus Medical Systems, Tokyo, Japan). Bowel preparation quality was reported according to a standardized scale. <i>ACN was defined as an invasive carcinoma or advanced adenoma (≥ 1 cm in size, villous component, or high-grade dysplasia).</i>
Statistical approach	Multivariable logistic regression analysis with a stepwise selection procedure based on Akaike information criterion.
Limitations	<ol style="list-style-type: none"> (1) The study was conducted in a single center in Korea, and nearly all subjects were of Korean ethnicity, thus limiting the generalizability of the model to other ethnic groups. (2) Most of the individuals in the cohort were employed, and these populations may be healthier than unemployed populations, resulting to potential selection bias. (3) Limited measures used for alcohol consumption, only measuring the frequency of alcohol use without specification regarding the quantity, duration, and intensity of intake. (4) Sessile serrated adenomas with dysplasia were not identified as ACN unless they were larger than 1 cm in size or contained components of high-grade dysplasia or carcinoma. <p>Other limitations identified by the reviewers:</p> <ol style="list-style-type: none"> (5) Use of univariate analysis to identify potential predictors may cause bias in selection of potentially important predictors due to confounding factors. (6) Use of arbitrary risk threshold determination may be inefficient, raising concern about its potential clinical utility. (7) Expected/observed ratio, and other estimates that were important in the interpretability and transferability of the risk prediction models were not reported.
New findings	The scoring model based on both clinical and laboratory risk factors included age, sex, family history of colorectal cancer, smoking, body mass index, serum levels of fasting glucose, low-density lipoprotein cholesterol, and carcinoembryonic antigen. It showed superior discriminatory accuracy compared with APCS scoring system.

Yen 2014

Primary objective	To assess how much of the variation in incidence of colorectal neoplasia is explained by baseline fecal hemoglobin concentration and assess the additional predictive value of conventional risk factors.
Study location(s)	Keelung community-based integrated screening (KCIS) in Keelung City for model derivation and Changhua and Tainan, Taiwan, for external validation.
Recruitment period	2001 to 2007.
Population	Cohorts aged ≥ 40 years, who were invited to population-based screening for colorectal neoplasia, using FIT as part of KCIS program.
Exclusion criteria	Not reported.
Predictor measures	Self-administered structured questionnaire provided at the screening site was used to assess the lifestyle factors, dietary habits, family history of CRC, personal history of cancer, type 2 diabetes, hypertension, cerebrovascular disease, cardiovascular disease, etc..
Outcome measures	Used FIT as a periodical screening tool with a one-year inter-screening interval and fHb concentration ≥ 100 ng/mL as the cutoff for further clinical investigation. Ans were identified via colonoscopy following referral from the KCIS (screen-detected) or vial linkage to the Cancer Registry (clinically detected).
Statistical approach	Accelerated-failure time regression model.

Limitations	(1) Information on gene susceptibility and FHbC within the KCIS cohort is not available; thus, it is not possible to investigate whether FHbC is a surrogate endpoint in its own right or a modified factor (interaction) of these genetic markers.
New findings	The high predictive ability supported by a dose-dependent relationship between baseline FHbC and the risk of developing colorectal neoplasia suggests that FHbC may be useful for identifying vases requiring closer postdiagnosis clinical surveillance, as well as being an early indicator of colorectal neoplasia risk in general population.

Yeoh 2011

Primary objective	To develop and validate a clinical risk score predictive of risk for colorectal advanced neoplasia for Asian population.
Study location(s)	11 Asian cities, namely Bangkok, Guangzhou, Hong Kong, Jakarta, Kuala Lumpur, Manila, New Delhi, Seoul, Singapore, Taipei, and Tokyo.
Recruitment period	July 2006 to December 2007.
Population and sample size	Asian cohorts eligible for CRC screening colonoscopy. Derivation cohort: 860 participants. Development cohort: 1892 participants.
Exclusion criteria	Not reported.
Predictor measures	Only mentioned the use of questionnaire administered at the time of colonoscopy to capture the clinical and lifestyle information.
Outcome measures	Colonoscopy was used to detect advanced neoplasia. <i>Advanced colorectal neoplasia was defined as invasive carcinoma or advanced adenoma.</i>
Statistical approach	Multiple logistic regression with split samples to derive and internally validate the model.
Limitations	Not reported Other limitations observed by the reviewer: <ul style="list-style-type: none"> - Inappropriate handling of continuous predictors such as age, which was dichotomized upon the development of the model. - Use of split-sampling technique to internally validate the model, which could result in loss of valuable information due to reduced sample size in the derivation cohort. - Non-reporting of predictor measures, and non-blinding of outcome assessors. - Non-reporting of important parameters such as model calibration, specificity, sensitivity, net reclassification index, and net benefit of the model. - Insufficient identification of plausible risk threshold (e.g., risk cutoffs were arbitrarily identified).
New findings	The new proposed Asia-Pacific Colorectal Screening (APCS) score enables risk stratification using elementary clinical information on age, gender, family history, and smoking. Moreover, the APCS score can successfully predict the risk of colorectal advanced neoplasia in asymptomatic Asian subjects. High-risk groups have a four-fold higher risk compared with the average risk group.

Appendix 4: Risk of bias assessment of all 37 studies

Table S1. Detailed appraisal and judgements for the risk-of-bias assessment using PROBAST.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Arnau-Collell 2022	Unclear risk	Unclear risk	Low risk	High risk	Some concerns	Low concern	Low concern	High risk	Some concern
	Although appropriate data sources and eligibility criteria were reported, the reasons for excluding other participants were unclear due to the secondary nature of the data source.	Insufficient information about the blinding of experts to predictor assessment.	Model outcome well defined and was based on WHO criteria.	Only predefined predictors were collected (age, sex, and FIT value). Other influential predictors were not considered in the analysis.	Limited to participants with FIT-positive results only, which limited the generalizability of the model.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	At least one domain has an unclear risk of bias, and all other domains have a low risk.	Some concerns in at least one domain, and low concerns in other domains.
Auge 2014	High risk	Low risk	Low risk	Unclear risk	Low concern	Low concern	Low concern	High risk	Low concern
	Potential selection bias due to inclusion of only FIT-positive individuals, predominantly comprising Caucasians.	Predictor variables were well defined, were assessed using a standardized tool, and were collected before the outcome was observed.	Model outcome well defined and standard consistent for all participants.	No sufficient information to make judgments regarding the appropriateness of handling missing data.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	At least one domain has an unclear risk of bias, and all other domains have a low risk.	Low concerns for all domains.
Brand 2017	Low risk	Low risk	Unclear risk	Unclear risk	Low concern	Low concern	Some concerns	Unclear risk	Some concern
	Appropriate data sources and eligibility criteria were reported.	Predictor variables were well defined. were assessed using a standardized tool, and were collected before the outcome was observed.	Outcome definition missing detailed definition of how colorectal adenoma was detected.	It was not clear whether there was a presence of competing risk. The number of iterations used during multiple imputation is insufficient.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition is unique to the rest of the included studies, making it hard to judge its applicability, but is still relevant to the review question.	At least one domain has an unclear risk, but the rest have low risk.	Some concerns in at least one domain, and low concerns in other domains.
Brigge 2022	Unclear risk	Low risk	Unclear risk	Unclear risk	Low concern	Low concern	Low concern	Unclear risk	Low concern
	Although appropriate data sources and eligibility criteria were reported, the reasons for	Predictor variables were well defined, were assessed using a standardized tool, and were	Outcome definition in the report missed a detailed definition of how	Use of complete case analysis by excluding participants with incomplete QCancer-10 colorectal cancer predictor data.	Included participants matched with the review question.	Definition, assessment, or timing of predictors	Outcome definition, timing, or determination	At least one domain has an unclear risk, but	Low concerns for all domains.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Cai 2012	excluding other participants were unclear due to the secondary nature of the data source.	collected before the outcome was observed.	colorectal adenoma was detected.			matched with the review question.	matched the review question.	the rest have low risk.	
	Unclear risk	Unclear risk	Low risk	Unclear risk	Low concern	Low concern	Low concern	Unclear risk	Low concern
	Potential selection bias observed in the validation cohort due to some hospitals having preselected groups with higher socioeconomic status.	Insufficient information about the blinding of experts to predictor assessment.	Model outcome was well defined, and standard consistent for all participants.	Insufficient information regarding how missing data were handled and lack of E/O estimates.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	Most of the domains have an unclear risk of bias.	Low concerns for all domains.
Cao 2015	Low risk	Low risk	Low risk	Unclear risk	Low concern	Low concern	Low concern	Unclear risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	The predictors were handled in the same way for all participants and assessed without knowledge of outcome data.	Model outcome was well defined, and standard was consistent for all participants.	The information provided is unclear regarding whether complexities in the data (e.g., censoring and competing risk), were appropriately accounted for.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	One domain has an unclear risk of bias, and the rest have a low risk.	Low concerns for all domains.
Chen 2014	Low risk	Low risk	Low risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	The predictors were handled in the same way for all participants and assessed without knowledge of outcome data.	Model outcome was well defined, and standard was consistent for all participants.	Use of univariate analysis to determine potential predictors. Inappropriate handling of continuous predictors. No detailed information about how missing data were handled; there was only some information about data exclusion.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	At least one domain has a high risk of bias.	Low concerns for all domains.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Cooper 2020	Low risk	Low risk	Unclear risk	Unclear risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	Potential transcription errors were minimized as clinical list developed were subject to a double reviewing process for code set validation.	The diagnosis and definition of CRC or polyps were not provided.	Used univariable Cox regression to assess 28 clinical features; insufficient information regarding handling of competing risk.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	At least one domain has an unclear risk of bias, and all other domains have a low risk.	Low concerns for all domains.
Deng 2023	High risk	High risk	Unclear risk	Unclear risk	Low concern	Low concern	Low concern	High risk	Low concern
	Potential selection bias due to nonrandom nature of selection of participants for confounders. Controls were selected based on a negative colonoscopy, and matching was limited only to sex and age.	Due to the retrospective nature of the study, predictors were likely measured after the outcome was known.	No sufficient information regarding blinding of outcome assessors to the predictor.	No sufficient information on how missing data were handled.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	At least one domain has unclear risk of bias, and with all other domains have a low risk.	Low concerns for all domains.
He 2019	Low risk	Unclear risk	Low risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	Predictors were assessed using standardized measures, but it was unclear whether their assessment was made without the knowledge of outcome data.	Used standardized colonoscopy conducted by expert endoscopist in a double-blinded approach in all study site. Clear time interval of predictor and outcome measurement was observed.	Continuous predictors such as age and BMI were converted to dichotomous predictors, which could result in loss of valuable information.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question	Outcome definition, timing, or determination matched the review question.	At least 1 domain has a high risk of bias.	

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Hong 2017	Low risk	High risk	Low risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Data sources and eligibility criteria were reported.	The predictors were defined and assessed equally in patients and were available when is going to be used, but the blinding of predictors was not possible.	The outcome was diagnosed using a standard measure without knowledge of predictors in all participants.	Participants with missing data were excluded, and the study used univariate over multivariate analysis to include predictors in the model.	Included participants highly matched with the review question.	Definition, assessment, and timing of predictor matched with review question.	Definition, assessment, and timing of outcome did match completely with review question.	At least one domain has a high risk of bias.	Low concerns for all domains.
Hyun Kim 2015	Low risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Data sources and eligibility criteria were reported.	Unclear definition of predictors and no sufficient information regarding blinding of assessors.	No sufficient information with regard to the blinding of outcome assessors.	The study used univariate over multivariate analysis to include predictors in the model, and there was inappropriate handling of continuous variables. Used split-sampling approach to internally validate the model.	Included participants highly matched with the review question.	Definition, assessment, and timing of predictor matched with review question.	Definition, assessment, and timing of the outcome matched with review question.	At least one domain has a high risk of bias	Low concerns for all domains.
Imperiale 2015	Low risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were well reported.	Predictor variables were well defined but were assessed using a self-administered questionnaire. It is unclear whether assessors were blinded to predictor assessment.	The colonoscopy and pathology results were obtained, reviewed, and coded by a trained study investigator. However, the AN definition was unclear.	The study used univariate over multivariate analysis to include predictors in the model. Used split sampling approach to internally validate the model.	Included participants highly matched with the review question.	Definition, assessment, and timing of predictor matched with review question,	Definition, assessment, and timing of outcome did match completely with review question	At least one domain has a high risk of bias	Low concerns for all domains.
Imperiale 2021	Low risk	Low risk	Low risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were well reported.	Predictor variables were well defined and assessed using a 50-item questionnaire.	The colonoscopy and pathology results were obtained, reviewed, and coded by a trained study investigator who was blinded to the survey/predictors. The outcome was well defined	Use of split-sample technique in validating the derived model. Also, there is insufficient information with regard to how competing risk was addressed nor how censoring was performed.	Included participants highly matched with the review question.	Definition, assessment, and timing of predictor matched with review question.	Definition, assessment, and timing of outcome did match completely with review question.	At least one domain has a high risk of bias.	Low concerns for all domains.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
			and standard consistent for all participants.						
Jung 2017	High risk	Low risk	Unclear risk	High risk	High concern	Low concern	Low concern	High risk	High concern
	Exclusion of participants due to incomplete data could introduce selection bias.	Predictor variables were well defined. Were assessed using a standardized tool, and were collected before the outcome was observed.	No sufficient information on whether outcome assessors were blinded to the predictors.	Selection of potential predictors were based on a univariate logistic regression analysis. Complete case analysis was also used to handle missing data.	Exclusion of individuals due to incomplete data could decrease the generalizability of the findings.	Definition, assessment, and timing of predictor matched with review question.	Definition, assessment, and timing of outcome did match completely with review question.	At least one domain has a high risk of bias.	At least one domain with high concern.
Jung 2018	High risk	Low risk	Low risk	High risk	High concern	Low concern	Low concern	High risk	High concern
	Exclusion of participants due to incomplete data could introduce selection bias.	Predictor variables were well defined. Were assessed using a standardized tool, and were collected before the outcome was observed.	The outcome was diagnosed using a standard measure without knowledge of predictors in all participants.	Participants with missing data were excluded, and the study used univariate over multivariate analysis to include predictors in the model, only Hosmer–Lemeshow test was used to evaluate calibration, and regression coefficients in the final model did not correspond to results from the multivariable analysis.	Exclusion of individuals due to incomplete data and inclusion of only FIT-negative participants could decrease the generalizability of the findings.	Definition, assessment, and timing of predictor matched with review question.	Definition, assessment, and timing of outcome did match completely with review question	At least one domain has a high risk of bias.	At least one domain with high concern.
Kaminski 2014	Low risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	Potential recall bias and selection bias could be observed due to the use of self-reported data to assess predictors.	Unclear risk due to no information regarding the non-blinding of outcome assessors.	Insufficient information on how missing data were handled. Continuous predictors were dichotomized, leading to potential loss of valuable information.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	At least one domain with a high risk of bias.	Low concerns for all domains.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Kim 2019	Low risk	High risk	Low risk	High risk	Low concern	High risk	Low concern	High risk	High concern
	Appropriate data sources and eligibility criteria were reported.	Cross-sectional collection of prediction could lead to misinterpretation of risk factors, leading to incorrect estimates of the model to predict the risk of CRC.	The outcome was diagnosed using a standard measure without knowledge of predictors in all participants.	Insufficient information on how missing data were handled and use of univariable regression analysis to determine potential predictors.	Included participants matched with the review question.	Potential measurement error and temporal bias can be observed due to timing of predictor assessment.	Outcome definition, timing, or determination matched the review question.	At least one domain has a high risk of bias.	At least one domain with high concern.
Liu 2018	High risk	Low risk	Unclear risk	High risk	Low concern	Low concern	High concern	High risk	High concern
	Exclusion of participants due to missing data may cause selection bias.	Predictors were defined and assessed in a similar way for all participants. Self-administered questionnaires were also validated in the target population.	No sufficient information regarding the timing of outcome assessment.	The exclusion of participants who have missing data in at least one of the risk factors of colon cancer could lead to potential selection bias.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	More focused on comparing different statistical approaches to risk prediction modeling.	At least one domain has a high risk of bias.	Low concerns for all domains.
Luu 2021	High risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Exclusion of participants due to missing data may cause selection bias.	Predictors were obtained using a structured questionnaire by an interviewer, but it is unclear whether assessors were blinded to the outcome.	No sufficient information regarding the timing of outcome assessment and whether assessors were blinded to the predictors.	Used univariate analysis to determine potential predictors. Age predictor was categorized. Excluded data with incomplete information.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	Low concerns for all domains.
Ma 2010	High risk	Unclear risk	Low risk	High risk	High concern	Low concern	Low concern	High risk	High concern
	Exclusion of people with competing risk, e.g., who have cardiovascular disease and diabetes, and those who have family history of CRC were also excluded.	No sufficient information regarding the use of validated questionnaire for physical activity and alcohol consumption predictors.	Model outcome well defined and standard consistent for all participants.	Use of univariate analysis to determine potential predictors. Exclusion of people with cardiovascular disease. Used Cox regression, but researchers did not account for competing risk.	Relevant target groups were excluded, which could affect the relevance of the model to the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	At least one domain with high concern.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Meester 2022	Unclear risk	Low risk	Unclear risk	Unclear risk	Low concern	Low concern	Low concern	Unclear risk	Low concern
	Data collected for ground truth from those who were positive according to the FIT test.	The predictors were handled in the same way for all participants and assessed without knowledge of outcome data.	There was no specification indicating if the outcome was blinded in the process.	Although predictors were determined using multivariate logistic regression, there is no sufficient information of how missing data were handled in the analysis.	Included participants matched the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has unclear risk, and the rest have low risk of bias.	Low concerns for all domains.
Müdler 2023	Unclear risk	Low risk	Unclear risk	Unclear risk	Low concern	Low concern	Low concern	High risk	Low concern
	Data collected for ground truth from those who were positive according to the FIT test.	The predictors were handled in the same way for all participants and assessed without knowledge of outcome data.	There was no specification indicating if the outcome was blinded in the process.	Although predictors were determined using multivariate logistic regression, there is no sufficient information of how missing data were handled in the analysis.	Included participants matched the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has unclear risk, and the rest have low risk of bias.	Low concerns for all domains.
Murchie 2017	High risk	Unclear risk	Unclear risk	Unclear risk	High concern	Low concern	Low concern	High risk	High concern
	Exclusion of people with competing risk or comorbidities and those who have family history of CRC.	No sufficient information regarding the use of validated questionnaire and blinding of the outcome.	No sufficient information regarding the timing of outcome assessment and whether assessors were blinded to the predictors.	No sufficient information on how competing risk or censoring bias was addressed.	Relevant target groups were excluded, which could affect the relevance of the model to the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	At least one domain with high concern.
Musselwhite 2019	High risk	Low risk	Low risk	High risk	High concern	Low concern	Low concern	High risk	High concern
	Exclusion of women and observations with missing data pose a serious concern regarding selection bias.	Predictors assessed the same across participants and predictors assessed without knowledge of outcome data.	The outcome was well defined and standard consistent across participants; the time interval between predictors and outcome measurement was appropriate.	Complexities in the data such as competing risk not properly accounted for and missing data not handled appropriately (they excluded).	Exclusion of women due to limited sample size and missing data.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	At least one domain with high concern.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Park 2019	High risk	High risk	Unclear risk	High risk	High concern	Low concern	Low concern	High risk	High concern
	Exclusion of people who have family history of CRC, which would limit the generalizability of the model to this population.	No blinding of the outcome. The family history of CRC was also mentioned as part of the predictors; this is confusing because it also was mentioned in their exclusion criteria.	No sufficient information regarding the timing of outcome assessment and whether assessors were blinded to the predictors.	Use of univariate analysis to determine potential predictors for CRC. Information is insufficient to make judgement particularly model overfitting, underfitting, and optimism.	Relevant target groups were excluded, which could affect the relevance of the model to the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	At least one domain with high concern.
Ruco 2015	Unclear risk	High risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Potential selection bias, as participants who provided informed consent might be significantly different from those who did not provide the consent.	Non-validated questionnaires were used to collect information on risk factors.	No sufficient information regarding the timing of outcome assessment and whether assessors were blinded to the predictors.	Use of univariate analysis to determine potential predictors for CRC. Age was presented in a categorical variable.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	Low concerns for all domains.
Schroy III 2015	Unclear risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Potential selection bias, as participants who have family history of CRC were excluded.	No sufficient information regarding blinding of assessors to outcome.	No sufficient information regarding the timing of outcome assessment and whether endoscopist was blinded to the survey.	Use of univariate analysis to determine potential predictors for CRC. Age was presented in a categorical variable.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	Low concerns for all domains.
Sekiguchi 2018	Low risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	No sufficient information regarding blinding of assessors to outcome.	No sufficient information regarding the timing of outcome assessment and whether endoscopist was blinded to the survey.	Use of univariate analysis to determine potential predictors for CRC. Age was presented in a categorical variable.	Included participants matched the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	Low concerns for all domains.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Sharara 2020	Low risk	High risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	Predictors were defined and assessed in a similar way for all participants, but researchers were not blinded to the outcome.	No sufficient information regarding the timing of outcome assessment and whether endoscopist were blinded to the survey.	Use of univariate analysis to determine potential predictors for CRC. Age was presented in a categorical variable. No information regarding how missing data were handled.	Included participants matched the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has a high risk of bias.	Low concerns for all domains.
Shin 2014	Low risk	Unclear risk	Low risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	It is unclear whether predictors were assessed the same across participants and whether predictors were assessed without knowledge of outcome data.	Outcomes were well-defined, assessed using standard measures, and consistent across participants; the time interval between predictors was appropriate.	Complexities in the data such as competing risk not properly accounted for and missing data not handled appropriately (e.g., use of complete case analysis/exclusion of all observations with incomplete data).	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination match the review question.	At least one domain has a high risk of bias.	Low concerns for all domains.
Soonklang 2021	Low risk	Unclear risk	Unclear risk	High risk	Low concern	Some concern	Some concerns	High risk	Some concerns
	Appropriate data sources and eligibility criteria were reported.	It is unclear whether predictors were assessed the same across participants and whether predictors were assessed without knowledge of outcome data.	No sufficient information regarding the timing of outcome assessment and whether endoscopist were blinded to the survey.	Did not perform sample size estimation (potentially insufficient sample size), and competing risk was not accounted for.	Included participants matched with the review question.	Definition, assessment, or timing of predictors were unclearly reported.	No sufficient information to make judgement regarding the timing of outcome assessment.	At least one domain has a high risk of bias.	Insufficient information to make judgement in most of the domains.
Sung 2018	High risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Exclusion of individuals with medical conditions. This could lead to selection, outcome, and ethical biases, especially for disadvantaged groups.	Predictors were assessed using standardized measures, but it was unclear whether their assessment was made without the knowledge of outcome data.	Used standardized colonoscopy conducted by expert endoscopist, but it was unclear whether experts who assessed the outcomes were blinded.	Continuous predictors such as age and BMI were converted to dichotomous predictors, which could result in loss of valuable information.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination matched the review question.	At least one domain has a high risk of bias.	Low concerns for all domains.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Stegeman 2014	Low risk	Low risk	Unclear risk	Low risk	Low concern	Low concern	Low concern	Unclear risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	Predictors were defined and assessed in a similar way for all participants and were assessed appropriately.	The outcome is well defined and standard consistent across participants, but it is unclear whether assessors were blinded to the predictors.	All statistical approaches for model development and validation were well considered and presented. Multiple imputation was used for dealing with missing data.	Included participants matched the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain with unclear risk and the rest has low risk.	Low concerns for all domains.
Sutherland 2021	Low risk	Low risk	Unclear risk	Low risk	Low concern	Low concern	Low concern	Unclear risk	Low concern
	Appropriate data sources and eligibility criteria were reported.	Predictors were defined and assessed in a similar way for all participants and were assessed appropriately.	The outcome is well defined and standard consistent across participants, but it is unclear whether assessors were blinded to the predictors.	All statistical approaches for model development and validation were well considered and presented. All necessary estimates were adequately reported. Multiple plausible risk thresholds based on predictive probabilities were used.	Included participants matched the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain with unclear risk and the rest has low risk.	Low concerns for all domains.
Tao 2014	High risk	Low risk	Low risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Exclusion of participants due to incomplete data on risk factors may cause selection bias.	Predictors were defined and assessed in a similar way for all participants and were assessed appropriately.	Outcomes were well defined, assessed using standard measures, and consistent across participants; the time interval between predictors was appropriate.	Continuous predictors such as age and BMI were converted to dichotomous predictors, which could result in loss of valuable information. Used split sample to internally validate the model and complete case analysis to handle missing data.	Included participants matched the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	At least one domain has high risk of bias.	Low concerns for all domains.
Thomsen 2022	Low risk	Low risk	Low risk	Unclear risk	High concern	Low concern	Low concern	Unclear risk	High concern
	Appropriate data sources and eligibility criteria were reported.	Predictors were defined and assessed in a similar way for all participants.	Authors reported the outcome in a well-defined manner and used standard/validate measures to assess the outcome.	No sufficient information regarding how missing data were handled.	Included participants included only persons who are FIT-positive.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately using standard methods.	At least one domain with unclear risk and the rest has low risk.	At least one of the domains has high concern.

Study ID	Risk of Bias				Applicability			Overall Judgement	
	Reasons for the Judgement				Reasons for the Judgement				
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	RoB	Applicability
Van 't Klooster 2020	Low risk	Low risk	Low risk	Low risk	High concern	Low concern	Low concern	Low risk	High concern
	Appropriate data sources and eligibility criteria were reported.	Predictors were defined and assessed in a similar way for all participants. Assessors were blinded to the outcome and to the study.	Authors reported the outcome in a well-defined manner and used standard/validate measures to assess the outcome.	All statistical approaches for model development and validation were well considered and presented. Competing risks were appropriately accounted for. All necessary estimates were adequately reported.	Included participants were limited to a population with cardiovascular disease only.	Definition, assessment, or timing of predictors matched with the review question.	Outcome was determined appropriately, using standard methods.	All domains have low risk of bias.	At least one of the domains has high concern.
Yang 2016	Low risk	Low risk	Low risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Used appropriate data source: cross-sectional and inclusion and exclusion criteria presented.	Predictors assessed the same across participants, and predictors assessed without knowledge of outcome data.	The outcome was well defined and standard consistent across participants. The time interval between predictors and outcome measurement was appropriate.	Missing data not handled appropriately by conducted complete case analysis. Used univariate analysis to determine potential predictors for model development. Competing risks were not accounted for.	Included participants matched with the review question.	Definition, assessment, or timing of predictors matched with the review question.	Outcome definition, timing, or determination match the review question.	At least one domain has high risk of bias.	Low concerns for all domains.
Yen 2014	Unclear risk	Unclear risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Unclear inclusion and exclusion criteria.	Predictors assessed the same across participants and predictors assessed without knowledge of outcome data.	The outcome was well defined and standard consistent across participants, but it is unclear whether assessors were blinded to the predictors.	Competing risk and “censoring” bias was not accounted for. No information regarding how missing data were handled.	Included participants matched with the review question.	Definition, assessment, or timing of predictors match with the review question.	Outcome definition, timing, or determination match the review question.	At least one domain has high risk of bias.	All domains showed low concerns regarding applicability.
Yeoh 2011	Low risk	High risk	Unclear risk	High risk	Low concern	Low concern	Low concern	High risk	Low concern
	Used appropriate data source: cross-sectional and inclusion and exclusion criteria presented.	Predictors were not well defined, and how they were assessed was unclearly reported.	Outcome definition and assessment were clearly reported, but blinding of outcome assessment was not reported.	Model derivation and validation used split sampling and handling of missing data were not reported.	Included participants matched with the review question.	Definition, assessment, or timing of predictors match with the review question.	Outcome definition, timing, or determination match the review question.	At least one domain has high risk of bias.	All domains showed low concerns regarding applicability.

Appendix 5. Evaluation of the potential clinical utility of all included risk prediction models for colorectal cancer

Table S2. Model performance metrics of risk prediction models for colorectal cancer and their projected clinical utility.

Study ID	Sample size (n)		Prevalence of AN (%)	Expected/ Observed Ratio	Sensitivity (%)	Specificity (%)	Handling of Continuous Predictors †	Risk Threshold Determination	Other Estimates Demonstrating Clinical Utility	Potential Clinical Utility
	DC (n)	VC (n)								
<i>Asia–Pacific Cancer Screening (APCS) risk prediction model</i>										
Yeoh 2011 ⁺⁺	860	1892	4.5 ¹ 3.0 ²	-	-	-	Age was categorized.	Three tiers (average, moderate, and high) of risk were arbitrarily defined.	-	⊕
Aniwan 2015 ⁺⁺⁺										
Chiu 2016 ⁺⁺⁺										
Sung 2018 ⁺⁺⁺	3829	1915	5.4 ¹ 6.0 ²	-	-	-	Age and BMI predictors were categorized.	Prevalence as threshold.	NNS _{AN} : <i>n</i> =18 NNR _C : <i>n</i> =11	⊕⊕
He 2019 ⁺⁺⁺	995	1201	4.1 ¹ 3.7 ²	-	76.7 ^C 36.7 ^{FIT}	-	Age and BMI predictors were categorized.	Artificially defined two risk tiers based on the AUC, Youden’s index, and Euclidian’s index.	NPV ^C : 98% NPV ^{FIT} : 97%	⊕⊕
Luu 2021 ⁺⁺⁺	12,520	-	2.5	-	48.0 ^{APCS} 68.0 ^{SG}	70.0 ^{APCS} 49.0 ^{SG}	Age was categorized.	Used similar risk threshold as the original APCS score by Yeoh et.al., 2011.	NPV ^{APCS} : 98%, PPV ^{APCS} : 4% NPV ^{SG} : 98%, PPV ^{SG} : 3%	⊕⊕⊕
Sekiguchi 2018 ⁺⁺⁺	5218	§	4.3	-	-	-	Age predictor (continuous) was transformed to categorical variable.	Threshold minimizing misclassification. The value minimizing the sum of the sensitivity and specificity was set as the risk threshold.	-	⊕⊕⊕
<i>Kaminski’s risk prediction model</i>										
Kaminski 2013 ⁺⁺	17, 979	17, 939	7.1	1.0	-	-	Age and BMI predictors were dichotomized, with age as continuous showing comparable c-statistic in the sensitivity analysis.	Prevalence as threshold.	-	⊕⊕
Ruco 2015 ⁺⁺⁺	-	5137	6.8	-	Presented range of sensitivity	Presented range of specificity	Age and BMI predictors were dichotomized.	Prevalence as threshold.	NPV: 93.16% PPV: 5.88%	

Study ID	Sample size (n)		Prevalence of AN (%)	Expected/ Observed Ratio	Sensitivity (%)	Specificity (%)	Handling of Continuous Predictors †	Risk Threshold Determination	Other Estimates Demonstrating Clinical Utility	Potential Clinical Utility
	DC (n)	VC (n)								
					across 8 tiers.	across 8 tiers.	(only for high-risk group)			
Other risk prediction model with external validation										
Cai 2012 ⁺⁺⁺	5229	2312	6.4	-	80.3	51.2	Age was categorized and presented by decade intervals.	The risk threshold was arbitrarily defined.	NNSc: n= 16	⊕⊕⊕
Liu 2018 ⁺⁺⁺	103, 249	-	1.12	1.05 ^M 1.19 ^F	-	-	Comparison of continuous and categorized lifestyle and dietary predictors was performed.	Unclear.	NRI ^M : 4.1% NRI ^F : -6.5%	⊕⊕
Musselwhite 2019 ⁺⁺⁺	3121	-	11.0	-	-	-	All continuous predictors were handled appropriately.	Based on 5-, 10-, and 20-year absolute risk.	-	⊕⊕
Shin 2014 ⁺⁺⁺	1,326,058	963,749	0.69	0.59-1.21 ^M 0.65-1.112 ^F	-	-	All continuous predictors were handled appropriately.	-	-	⊕
Thomsen 2022 ⁺⁺⁺	34,929	21,530	5.9	1.02	-	-	Age and FIT result predictors (continuous) were transformed to categorical variables.	Prevalence of FIT positive as threshold.	-	⊕⊕⊕⊕
Van 't Klooster 2020 ⁺⁺⁺	7280	9322	2.5	1.16 ^{CRC} 0.85 ^{CE}	-	-	All continuous predictors were handled appropriately.	NA	10-year predicted risk: 2% for both cohorts Absolute lifetime risk: 5% ⁺⁺ and 4% ⁺⁺⁺	⊕⊕⊕⊕
Yen 2014 ⁺⁺⁺	54,921	Unclear	-	-	-	-	BMI predictor (continuous) was transformed to categorical variable.	-	-	⊕⊕
De novo models without external validation										
Auge 2014 ⁺	3109	- -	9.5 9.5	- No	- No	-No	Age predictor (continuous) was transformed to categorical variables.	Risk threshold was arbitrarily defined.	-	⊕⊕
Brand 2017 ⁺⁺	9934	10,034	40●	1.01*	-	-	Appropriate handling of continuous predictors. Age and BMI were treated as continuous in the analysis.	-	-	⊕

Study ID	Sample size (n)		Prevalence of AN (%)	Expected/ Observed Ratio	Sensitivity (%)	Specificity (%)	Handling of Continuous Predictors †	Risk Threshold Determination	Other Estimates Demonstrating Clinical Utility	Potential Clinical Utility
	DC (n)	VC (n)								
Cao 2015 ++	17,970 ^W 4881 ^M	§	3.8 ^W 6.7 ^M	Reported E/O ratio for each decile (1-10 decile)	-	-	Compared model fit between continuous vs. categorical risk factor but used categorical risk factor for the final model.	Unclear.	-	⊕⊕
Chen 2014 ++	905	§	5.3	-	93.8	47.6	Age predictor (continuous) was transformed to categorical variables.	Unclear.	NPV: 99.3%, PPV: 9.1%, NNS ^{HR} : 11, NNS ^{LR} :137	⊕⊕
Cooper 2020 ++	292,059	§	5.41	1.0	58.82 ^C	91.38 ^C	All continuous predictors were handled appropriately.	Threshold minimizing misclassification was determined using the PPV and NPV and based on the NICE guideline.	PPV: 3% corresponds to a risk probability threshold of 0.0168	⊕⊕⊕
Hong 2017 ++	24,725	24,725	2.3	-	70.8	61.2	All continuous predictors were handled appropriately.	Prevalence as threshold.	NPV: 98.9% PPV: 4.0%	⊕⊕⊕
Hyun Kim 2015 ++	2152	1316	4.4	-	39.3 ^{HR vs. AR + LR}	85.8 ^{HR vs. AR + LR}	Age and BMI predictors (continuous) were transformed to categorical variables.	Risk thresholds were arbitrarily defined.	-	⊕⊕
Imperiale 2015 ++	2993	1467	9.4	-	-	-	All continuous predictors were handled appropriately.	Risk thresholds were arbitrarily defined.	-	⊕⊕
Imperiale 2021 ++	3025	1475	9.1	-	-	-	Age predictor (continuous) was transformed to categorical variable.	Prevalence as threshold.	-	⊕⊕
Jung 2017 ++	57,635	38,600	1.3	-	6.7.3	57.6	All continuous predictors were handled appropriately.	The risk threshold was only based on the Youden index, defined as sensitivity + specificity-1.	NNS: 50	⊕⊕
Jung 2018+	11,873 ^{FIT-}	-	2.1	-	27.0	89.0	BMI predictor (continuous) was transformed to categorical variable.	Risk thresholds were arbitrarily defined.	-	⊕
Kim 2019 ++	41,702	17,873	0.9	-	97.1 ^{HR vs. AR + LR}	10.7 ^{HR vs. AR + LR}	Age and BMI predictors (continuous) were dichotomized.	Risk thresholds were arbitrarily defined.	-	⊕

Study ID	Sample size (n)		Prevalence of AN (%)	Expected/ Observed Ratio	Sensitivity (%)	Specificity (%)	Handling of Continuous Predictors †	Risk Threshold Determination	Other Estimates Demonstrating Clinical Utility	Potential Clinical Utility
	DC (n)	VC (n)								
Ma 2010 ⁺⁺⁺	28,115	18,256	-	0.94	-	-	Age and BMI predictors (continuous) were categorized.	Used absolute 10-year risk probability and a reference standard by age group.	-	⊕⊕
Murchie 2017 ⁺⁺	5063	§	5.7	-	-	-	All continuous predictors were handled appropriately.	Unclear.	-	⊕
Park 2019 ⁺	3733	-	9.8	-	-	-	BMI predictor (continuous) was dichotomized using a standard cutoff for obesity.	Risk thresholds were arbitrarily defined.	-	⊕
Schroy III 2015 ⁺⁺	3543	§	5.7	-	-	-	Age and BMI predictors (continuous) were dichotomized.	Used predicted probability (prevalence) to determine risk threshold.	NRI _{smoking} : 8.4% NRI _{RES} : 2.2%	⊕⊕⊕
Sekiguchi 2018 ⁺⁺	5218	§	4.3	-	-	-	Age predictor (continuous) was transformed to categorical variable.	Threshold minimizing misclassification. The value minimizing the sum of the sensitivity and specificity was set as the risk threshold.	-	⊕⊕
Sharara 2020 ⁺⁺	980	§	5.10	-	-	-	BMI predictor (continuous) was dichotomized using a standard cutoff for obesity.	-	-	⊕
Soonklang 2021 ⁺⁺	1311	§	4.04	-	-	-	Age and BMI predictors (continuous) were transformed to categorical variables.	-	-	⊕
Stegeman 2014 ⁺	1121	-	9.1	-	40.0	93.0	Age and BMI predictors (continuous) were transformed to categorical variables.	Utility-based risk threshold determination.	NRI _{risk-based} : 5.4%	⊕⊕⊕
Sutherland 2021 ⁺⁺	3035	§	7.53	-	** 87.9, 52.6, 12.1, 2.81	** 33.4, 72.9, 97.2, 99.5	All continuous predictors were handled appropriately.	Presented multiple plausible risk threshold based on predicted probabilities.	NPV ^{**} : 96.9%, 94.5%, 92.5%, 92%.	⊕⊕⊕

Study ID	Sample size (n)		Prevalence of AN (%)	Expected/ Observed Ratio	Sensitivity (%)	Specificity (%)	Handling of Continuous Predictors †	Risk Threshold Determination	Other Estimates Demonstrating Clinical Utility	Potential Clinical Utility
	DC (n)	VC (n)								
									PPV**: 10.6, 14.8, 27.8, 31.8.	
Tao 2014 ++	7891	3519	9.9	-	-	-	Age and BMI predictors (continuous) were transformed to categorical variables.	Risk thresholds were arbitrarily defined.	NNS: 9	⊕⊕
Yang 2016 ++	49,130	21,052	1.4	-	-	-	Age and BMI predictors (continuous) were transformed to categorical variables.	Risk thresholds were arbitrarily defined.	-	⊕⊕

Symbols: -, not reported; + development study (without internal validation); ++ development study with internal validation; +++ (with) external validation study; ++++ model updating; ¹ derivation cohort; ² internal validation cohort; † in the context of individualized risk prediction model; • refers to the prevalence of any colorectal adenoma that are histologically confirmed; ^c combined model test results; ^{FIT}, FIT result only; ^w, women; ^M, men; §, used similar cohort by using 10-fold cross-validation or bootstrapping method; ^{HR}, high risk; ^{LR}, low risk; ^{AR}, average risk; ≈, based on cutoff of >2; ^{SG}, screening guideline; ^{RES}, inclusion of race/ethnicity by sex interaction; CRC, colorectal cancer; ^{CE}, competing event/risk; **, based on cutoffs of 5%, 10%, 20%, 30%, and 40%, respectively. **Abbreviations:** DC, derivation cohort; VC, validation cohort; NRI, Net Reclassification Index; NPV, negative predictive value; NNS_{AN}, number needed to screen for detecting one advanced neoplasia; NNRC, number needed to refer for colonoscopy; PLR, positive likelihood ratio; NRI^{M,F}, net reclassification index in men and in women, respectively. ⊕⊕⊕⊕, **High potential:** interpretable performance metrics, risk estimates, and (multiple) plausible risk thresholds were reported. Appropriate handling of continuous predictors and risk threshold determinations were observed. ⊕⊕⊕, **Some concerns:** Minor issues exist in terms of performance metrics reporting or risk threshold determination, including few (≤2) missing performance metrics, and use of arbitrary risk thresholds, respectively. ⊕⊕, **Low potential:** limited potential for clinical utility due to inappropriate handling of continuous predictors, unclear model calibration, risk threshold determination, sensitivity, specific, and other estimates demonstrating clinical utility. ⊕, **Very low potential:** very minimal information provided, making it hard to assess any meaningful clinical utility.

Appendix 6. Search strategies

MEDLINE (Through PubMed) Search date 24 November 2022	#1	("Colorectal Neoplasms"[Mesh] OR ((colorectal[Title/Abstract] AND (neoplasm[Title/Abstract] OR neoplasms[Title/Abstract] OR cancer[Title/Abstract] OR cancers[Title/Abstract] OR tumor[Title/Abstract] OR tumors[Title/Abstract] OR carcinoma[Title/Abstract] OR carcinomas[Title/Abstract])))	Search #1 AND #2 AND #3 AND #4 AND #5 Results: 1,529
	#2	("Risk"[Mesh] OR (risk[Title/Abstract] OR risks[Title/Abstract]))	
	#3	(Models OR model*[Title/Abstract])	
	#4	(predict*[Title/Abstract] OR assess*[Title/Abstract] OR estimat*[Title/Abstract/Abstract])	
	#5	(individual*[Title/Abstract] OR personal*[Title/Abstract] OR particular*[Title/Abstract])	
	#6	#1 AND #2 AND #3 AND #4 AND #5	
The Cochrane Library Search date 2 November 2022	#1	MeSH descriptor: [Colorectal Neoplasms] explode all trees	
	#2	colorectal:ti,ab,kw and cancer or cancers or neoplasm or neoplasms or tumor or tumors or carcinoma or carcinomas:ti,ab,kw (Word variations were searched)	
	#3	1 OR 2	
	#4	MeSH descriptor: [Risk] explode all trees	
	#5	risk or risks:ti,ab,kw (Word variations were searched)	
	#6	4 or 5	
	#7	MeSH descriptor: [Models, Theoretical] explode all trees	
	#8	model*:ti,ab,kw (Word variations were searched)	
	#9	7 or 8	
	#10	predict* or assess* or estimat*:ti,ab,kw (Word variations were searched)	
	#11	individual* or personal* or particular*:ti,ab,kw (Word variations were searched)	
	#12	3 and 6 and 9 and 10 and 11	
MEDLINE, Global Health, and Biological Abstracts	#1	exp Colorectal Neoplasm/	Search: 5 AND 8 AND 11 AND 12 AND 13 remove duplicates from 15
	#2	colorectal.ab,kf,ti.	
	#3	neoplasm.ab,kf,ti. OR neoplasms.ab,kf,ti. OR cancer.ab,kf,ti. OR cancers.ab,kf,ti. OR tumor.ab,kf,ti. OR tumors.ab,kf,ti. OR carcinoma.ab,kf,ti. OR carcinomas.ab,kf,ti	

(Through Ovid)	#4	2 AND 3	limit 16 to English language Results: 2093
	#5	1 OR 4	
	#6	exp Risk/	
	#7	risk.ab,kf,ti. OR risks.ab,kf,ti.	
	#8	6 OR 7	
	#9	exp theoretical model/	
	#10	model*.ab,kf,ti.	
	#11	9 OR 10	
	#12	predict*.ab,kf,ti. OR assess*.ab,kf,ti. OR estimat*.ab,kf,ti.	
	#13	individual*.ab,kf,ti. OR personal*.ab,kf,ti. OR particular*.ab,kf,ti.	
	#15	5 AND 8 AND 11 AND 12 AND 13	
	#16	remove duplicates from 15	
	#17	limit 16 to English language	
Web of Science	#1	"Colorectal Neoplasm" (Topic)	
	#2	Colorectal cancer OR colorectal carcinoma OR colorectal tumor* (Topic)	
	#3	risk OR risk factors (Topic)	
	#4	prediction model* (Topic)	
	#5	predict* OR estimate* OR assess* (Topic) AND individual* OR personal* OR particular* (All Fields)	
	#6	#1 OR #2 AND #3 AND #4 AND #5	