Heliyon 10 (2024) e33619

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

A self-reported symptom-based decision-making model helps to rule out outpatient cases at low risk for CRC before colonoscopy

Zhe Luan [a,1], Fangfang Liu [b,1], Li Zhang [c,1], Jun Chen [a], Yiming Zhao [d], Congyong Li [e], Zhaoyun Liu [f], Huawei Li [g], Li Dong [c], Funing Zang [h], Lingyan Han [i], Tianyue Zhao [g], Qiao Wang [f], Gang Sun [a], Shufang Wang [a,*]

[a] Department of Gastroenterology and Hepatology, The First Medical Center, Chinese PLA General Hospital, Beijing, 100853, China
[b] Department of Oncology, 5th Medical Center of Chinese PLA General Hospital, Beijing, 100853, China
[c] Department of Gastroenterology, Taiyuan Central Hospital, Shanxi Medical University, Taiyuan, 030009, China
[d] Department of Gastroenterology and Hepatology, Hainan Hospital of Chinese PLA General Hospital, Sanya, 572013, China
[e] Sixth Health Care Department, Second Medical Center of PLA General Hospital, Beijing, 100853, China
[f] Department of Gastroenterology, Ningde Hospital of Traditional Chinese Medicine Affiliated to Fujian University of Traditional Chinese Medicine, Ningde, 352100, China
[g] Department of Gastroenterology, Panjin Central Hospital, Panjin, 124013, China
[h] Department of Gastroenterology, Pizhou City Hospital of Traditional Chinese Medicine, Pizhou, 221300, China
[i] Department of Gastroenterology, Shanxi Traditional Chinese Medical Hospital, Taiyuan, 030012, China

ARTICLE INFO

Keywords:
Colorectal cancer
Decision-making model
Colonoscopy
Outpatient
Risk prediction
Nomogram

ABSTRACT

Objectives: Effective exclusion of low-risk symptomatic outpatient cases for colorectal cancer (CRC) remains diagnostic challenges. We aimed to develop a self-reported symptom-based decision-making model for application in outpatient scenarios.

Methods: In total, 8233 symptomatic cases at risk for CRC, as judged by outpatient physicians, were involved in this study at seven medical centers. A decision-making model was constructed using 60 self-reported symptom parameters collected from the questionnaire. Further internal and external validation cohorts were built to evaluate the discriminatory power of the CRC model. The discriminatory power of the CRC model was assessed by the C-index and calibration plot. After that, the clinical utility and user experience of the CRC model were evaluated.

Results: Nine symptom parameters were identified as valuable predictors used for modeling. Internal and external validation cohorts verified the adequate discriminatory power of the CRC model. In the clinical application step, all 17 physicians found the model easy to grasp, 99.9 % of the patients were satisfied with the survey form. Application of this model detected all CRC cases. The total consistency ratio of outpatient cases undergoing colonoscopy was 81.4 %. None of the low-risk patients defined by the CRC model had been diagnosed with CRC.

Conclusion: This multicenter study developed and validated a simple and user-friendly decision-making model covering self-reported information. The CRC model has been demonstrated to perform well in terms of rapid outpatient decision-making scenarios and clinical utility, particularly because it can better rule out low-risk outpatient cases.

* Corresponding author.
E-mail address: wangshufang2006@aliyun.com (S. Wang).
[1] ZL, FFL and LZ contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e33619
Received 9 December 2023; Received in revised form 23 June 2024; Accepted 24 June 2024
Available online 25 June 2024

1. Introduction

Colorectal cancer (CRC) is the third most common cancer globally and the second most common cancer in China [1,2]. China ranks among the top five worldwide with a prevalence rate of 37.6 per 100,000 and a 4.2 % rate of increase [3]. Colonoscopy is still the gold standard for early CRC detection, and it is now quite accepted that people over the age of 50 should undergo regular colonoscopy for cancer screening [4,5]. The coverage of colonoscopy screening in the United States for people aged 50 and above increased from 21 % in 2000 to 60 % by 2015 [5]. In developing countries, this strategy is limited by insufficient resources for endoscopy. Taking China as an example, the coverage rate of colonoscopy is 15 % [6]. The gap between people who need a colonoscopy and those who receive a colonoscopy is large. How to allocate and utilize colonoscopy resources reasonably and effectively is worth considering.

On the other hand, populations with or without digestive symptoms do not always need to be examined in the short term. The outpatient doctors prefer to recommend colonoscopy based on personal clinical experience, which tends to be more common in areas with less developed medical care [7]. From 2000 to 2016, we initiated a clinical retrospective study that enrolled 34,2922 patients who underwent colonoscopy in five regions of China, most of whom were patients with abdominal symptoms and considered at risk for CRC by their outpatient physicians. The results suggested that nearly 50 % of the patients showed no abnormality on colonoscopy, and the positive detection rate of the examined population was low [6].

To ensure the proper allocation of outpatient endoscopy resources, quantitative methods such as risk scores or decision-making models limit subjectivity. They can improve the detection rate of CRC, especially for areas with unequally distributed and limited health care resources. To date, a series of decision-making models have already been developed on request [8–13]. One of the most widely used model is The Asia-Pacific Colorectal Screening (APCS) scoring system. The APCS score is based on four factors: age, sex, family history of CRC, and smoking. In 2014, the improved APCS model was modified to add two variables, BMI and self-reported diabetes [14]. The improved APCS scoring system has been proposed and recommended for CRC risk assessment by domestic CRC screening experts, so as to concentrate on the high-risk groups. However, the improved APCS scoring system still has some shortcomings, and important symptom information loss is sacrificed at the expense of convenience and had relatively weak discriminatory power. Successful translation of these findings into specific scenarios, especially outpatient colonoscopy application, remains a major challenge [8,15–17].

The aim of this multicenter study was to develop a clinical decision-making model that covers essential symptom information and lifestyle risk factors for patients. The subjective evaluation of outpatient cases and physicians in the process of use, the ratio of clinical decision changes and the influence on the final diagnosis results were collected. We hypothesized that the self-reported symptom-based decision-making model would be efficient and cost-effective for outpatient doctors, and is not inferior to the APCS scoring

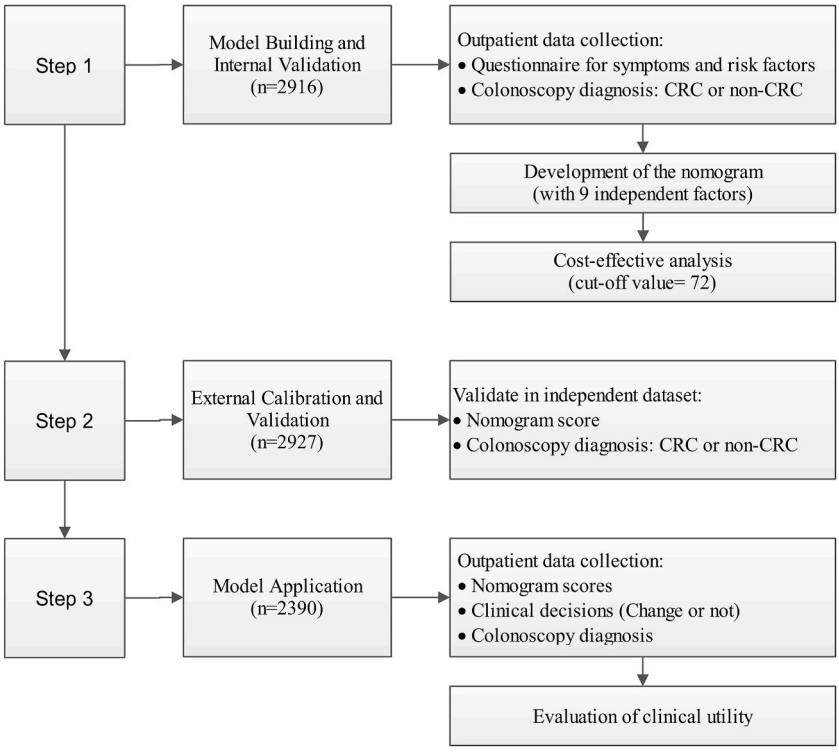

Fig. 1. The study flowchart. The flowchart depicts the design of our study, which consists of the following three steps: model building and internal validation, external validation and calibration, and model application. The model's discriminative ability was assessed by the C-index and calibration in all three steps. The clinical utility was evaluated in Step 3.

system.

2. Methods

2.1. Study design and participants

This observational study consisted of the following three steps: model building and internal validation (Step 1), external validation and calibration (Step 2) and model application (Step 3). The study protocol was approved by the Chinese PLA General Hospital Ethics Committee (301hn11-2017-05). Written informed consent was obtained from all patients for the publication of all their data and/or images. Qualification of the researcher and diagnostic criteria is shown in Supplemental Digital Content. The study flowchart is presented in Fig. 1. From 2017 to January 2022, we enrolled three batches of symptomatic outpatient cases from seven hospitals in China (The First Medical Center of Chinese PLA General Hospital, Hainan Hospital of Chinese PLA General Hospital, Taiyuan Central Hospital, Shanxi Traditional Chinese Medical Hospital, Ningde Hospital of Traditional Chinese Medicine Affiliated to Fujian University of Traditional Chinese Medicine, Panjin Central Hospital, Pizhou City Hospital of Traditional Chinese medicine). In the Step 1, we used the data of The First Medical Center of Chinese PLA General Hospital (n = 2916), and in the Step 2, we used the data of Hainan Hospital of Chinese PLA General Hospital (n = 2927). In the Step 3, we collected the subjective evaluation of outpatient doctors in the process of use, the ratio of clinical decision changes, and the influence on the final diagnosis results from another 5 hospitals (n = 2390). In total, 8233 eligible symptomatic outpatients were recruited from 7 hospitals in China.

The inclusion criteria were: (1) male/female, regardless of age; (2) symptomatic outpatients; (3) signing informed consent; (4) a previous colonoscopy with a diagnosis of related colorectal disease may be included. The exclusion criteria were: (1) any contraindications to colonoscopy; (2) asymptomatic individuals; (3) pregnancy; (4) cardiopulmonary insufficiency. All patients met our inclusion and exclusion criteria.

The data collected from Step 1 were used for model development and internal validation. A total of 2916 symptomatic patients who underwent colonoscopy were included in this step. They were asked to complete a questionnaire giving details of 60 symptoms and risk factors for their condition (more details see Supplementary Material). Questionnaire data were collected from all participants. Trained investigators helped participants complete questionnaires. All participants underwent colonoscopy examination. In the internal validation phase, the model was compared with the improved APCS scoring system. The improved APCS scoring system: six variables including age, gender, first-degree relative family history of CRC, BMI, self-reported diabetes and smoking were included. The scoring system divided the population into average-risk group (0–2 points) and high-risk group (3–6 points) according to the score.

After the model was tested and validated internally, a self-reported symptom-based nomogram for outpatient colonoscopy

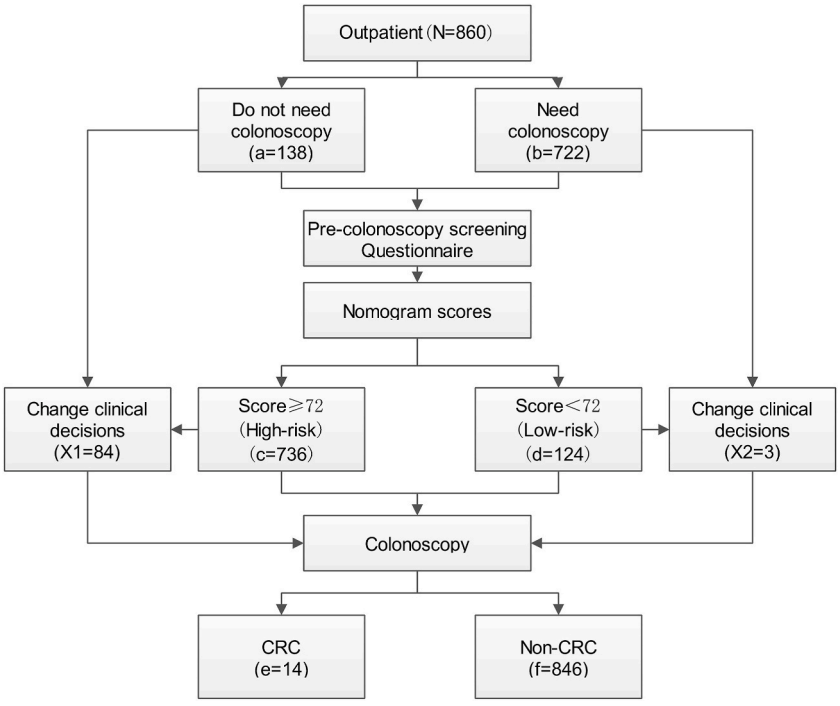

Fig. 2. Flowchart for Step 3: Model application. A total of 2390 outpatient cases were enrolled in this step. Secondary to the outpatient doctor making initial clinical decisions (a and b), nomogram scores were calculated from information collected in the outpatient department. After the model gave the decision suggestions (c and d), the physicians made the final decision, while at the same time, clinical decision changes were recorded (X1 and X2). Finally, all subjects considered for further examination underwent colonoscopy for CRC screening (e and f).

application was developed. The data collected from Step 2 (n = 2927) were used for external validation. In Step 3, the subjective evaluation of outpatient doctors in the process of use, the ratio of clinical decision changes, and the influence on the final diagnosis results were collected (Fig. 2).

2.2. Statistical analysis

No generally accepted approaches exist to estimate sample size requirements for derivation and validation studies of risk decision-making models. We used all available data to maximize the power and generalizability of our results. Model reliability was enhanced by our use of an external validation cohort. The outcome variable was categorized as a binary variable with CRC and non-CRC categories. Continuous variables are described as the mean ± standard deviation (SD). Categorical variables are described as frequencies and proportions. Student's *t-test* or the Mann–Whitney *U test* were used to compare continuous variables, and the χ^2 or Fisher's exact test were used to compare categorical variables as appropriate.

A multivariate logistic regression model was built using a stepwise selection of the variables collected from the questionnaire. Once the model was fitted, the scores from regression coefficients were generated for each patient corresponding to CRC probability. Odds ratios (ORs) with 95 % confidence intervals (CIs) were calculated to estimate the associations between various predictive variables and CRC probability. The area under the receiver operating characteristic curve (AUC) was employed to assess the accuracy of the models. The Hosmer–Lemeshow test was used to assess the goodness-of-fit of the logistic model. Optimal cutoff values were calculated by the maximum Youden indices. Stepwise selection was utilized as the method for feature selection. Through the application of this approach, we successfully narrowed down the initial 60 self-reported symptom parameters to nine predictors that most effectively encompassed the pertinent information for colorectal cancer screening.

A nomogram was developed based on the results of multivariate logistic regression analysis. The discrimination and calibration of the nomogram were assessed by Harrell's concordance index (C-index) and a calibration plot [18]. The C-index is equivalent to the area under the curve (AUC). An AUC of 0.6–0.7 was considered poor, 0.7–0.9 excellent, and >0.9 outstanding. A 1000-sample bootstrapped calibration plot was developed, which was considered to have good performance when the calibration curve closely resembled the line representing perfect calibration (the prespecified acceptable mean absolute error (MAE) for the calibration curve was <0.4).

All analyses used a two-sided P value of 0.05 as statistical significance. Statistical analyses were performed using SAS 9.4 (SAS Institute Inc. Cary. NC. USA) and R version 4.0.3 (https://rstudio.com).

Table 1
Characteristics of symptomatic patients who underwent colonoscopy.

Variables	Total		Non-CRC		CRC		*P*
	Number	%	Number	%	Number	%	
Age							
≤50	1415	48.53	1403	48.95	12	24.00	
>50	1501	51.47	1463	51.05	38	76.00	0.1113
Gender							
Male	1584	54.32	1560	54.43	24	48.00	
Female	1332	45.68	1306	45.57	26	52.00	<0.0001
[a]BMI							
≤25	281	9.64	278	9.70	3	6.00	
>25	2635	90.36	2588	90.30	47	94.00	<0.0001
Exercise	1197	41.05	1184	41.31	13	26.00	0.0291
Smoking	832	28.53	817	28.51	15	30.00	0.8167
Alcohol	978	33.54	964	33.64	14	28.00	0.4027
Redmeat	339	11.63	332	11.58	7	14.00	0.5972
Family CRC History	744	25.51	729	25.44	15	30.00	0.4630
Colonpolyps History	557	19.10	551	19.23	6	12.00	0.1976
CRC History	41	1.41	35	1.22	6	12.00	0.2633
Preserved Food	373	12.79	360	12.56	13	26.00	0.0048
Anemia	91	3.12	86	3.00	5	10.00	0.0048
Mucus bloody stool	317	10.87	295	10.29	22	44.00	<0.0001
[a]Ascites	14	0.48	11	0.38	3	6.00	<0.0001
Hypertension	487	16.70	469	16.36	18	36.00	0.0002
[a]Diabetes	219	7.51	217	7.57	2	4.00	0.3421
	Mean	SD	Mean	SD	Mean	SD	*P*
Age	53	12.95	53	12.90	62	12.62	0.8828
BMI	41	14.31	41	14.34	42	12.60	0.2501

CRC, Colorectal Cancer.
[a] Using Fisher's exact.

3. Results

3.1. Demographic characteristics of symptomatic patients

The demographic characteristics of the Step 1 population are listed in Table 1. A total of 2916 symptomatic patients were analyzed, and 50 (1.71 %) patients were diagnosed with CRC. Thirty-eight (76 %) patients diagnosed with CRC were older than 50 years, with a mean age of 62 ± 12.62 years. Twenty-six (52 %) patients were female. The reported protective factors in CRC cases were less than those in non-CRC cases, including exercise and colon polyp history ($P < 0.05$). More risk factors were reported, including smoking, alcohol, red meat, CRC history, family CRC history, preserved food, anemia, bloody-mucus stool, ascites, hypertension, and diabetes, than in non-CRC cases (all $P < 0.05$).

Differences in demographic and clinical variables were compared between CRC/non-CRC groups (Table 1). The preliminary comparison showed that there were significant differences in the CRC group and non-CRC group in the following aspects: sex, BMI, exercise level, drinking and smoking history, personal history of CRC, and family history, and significant differences in symptomatology with anemia (hemoglobin <100 g/L), bloody-mucus stool, ascites, hypertension and diabetes. The patients without a history of regular exercise (more than twice a week, on average, for any length of time) or a habit of eating preserved foods (more than twice a week, on average) were more prone to developing CRC, which was consistent with previous studies [19,20]. Risk factors for CRC were determined using a stepwise multiple logistic regression analysis, including the following: older age, preserved food, less exercise, anemia, ascites, bloody mucus stool, noncolon polyp history, CRC history, and hypertension were correlated with CRC.

3.2. Performance of the decision-making model

Table 2 summarizes the results of the multivariate regression analyses. Older age, preserved food, less exercise, anemia, ascites, bloody-mucus stool, colon polyp history, CRC history, and hypertension were independent predictors of CRC.

The model for CRC had an AU-ROC (95 % CI) of 0.8574 (0.8006–0.9142), sensitivity = 78.6 % and specificity = 79.2 %. The Hosmer–Lemeshow test showed the goodness-of-fit of this multivariate regression model ($P = 0.846$). We also compared the predictive ability of the CRC model to the improved Asia-Pacific scoring system (APCS) scoring system. The AU-ROC (95 % CI) of CRC was higher than the improved APCS of 0.6682 (0.5968–0.7396) ($P < 0.0001$) (Fig. 3(a), Table 3). Fig. 3(b) shows the calibration plot, suggesting a favorable agreement between the predictions and observations (MAE = 0.003).

After external validation using the bootstrap technique, the C-index of this nomogram was 0.8190 (95 % CI 0.7576–0.8795), sensitivity = 84.4 % and specificity = 69.6 % (Fig. 3(c)), which indicates adequate discriminatory power. Fig. 3(d) shows the calibration plot, suggesting a favorable agreement between the predictions and observations (MAE = 0.004).

3.3. Construction of the nomogram and cost-effective analysis

Fig. 4 shows the predictive nomogram derived from the β coefficients of the nine independent factors. The top row of the nomogram corresponds to the general score for each predictor listed on the left, and there was a corresponding row on the right indicating possible descriptors. After characterizing the patient for each predictor, a perpendicular line toward the first row should be drawn to identify the value. This action should be performed for all nine predictors and the final score tallied. This final score should be identified in a total point row, and then a perpendicular line is drawn that corresponds to the probability of CRC.

Table 4 describes the temporary cutoff values from the nomogram scoring system that were established to calculate the predictive performance at each cutoff. When the cutoff score was set at <19 for the predicted rate of CRC, none of the patients were diagnosed with CRC, while 145 patients underwent colonoscopy. When the cutoff was ≤72, 16(0.7 %) cases of CRC were observed, while 2286 underwent colonoscopy. When the cutoff was >72, 34 (5.4 %) cases of CRC were observed, while 630 underwent colonoscopy.

Table 2
Logistic model for predicting colorectal cancer.

Variables	OR	95%CI		*P*
		Lower	Upper	
Age	1.061	1.035	1.087	<0.0001
Preserved	2.163	1.073	4.364	0.0311
Exercise	0.391	0.197	0.776	0.0073
Anemia	3.317	1.175	9.366	0.0235
Ascites	13.014	3	56.442	0.0006
Mucus bloody stool	6.419	3.445	11.963	<0.0001
Colonpolyps history	0.243	0.088	0.67	0.0063
CRC history	7.671	2.373	24.8	0.0007
Hypertension	2.521	1.311	4.845	0.0056

CRC, Colorectal Cancer.

*, The corresponding logistic regression equation was as follows: logit (CRC) = −7.9694 + 0.0588 * age + 0.7717 * preserved food - 0.9392*exercise + 1.1992*anemia +2.5660*ascites + 1.8593*bloody-mucus stool – 1.4141*colon polyps + 2.0374*CRC + 0.9245*hypertension.

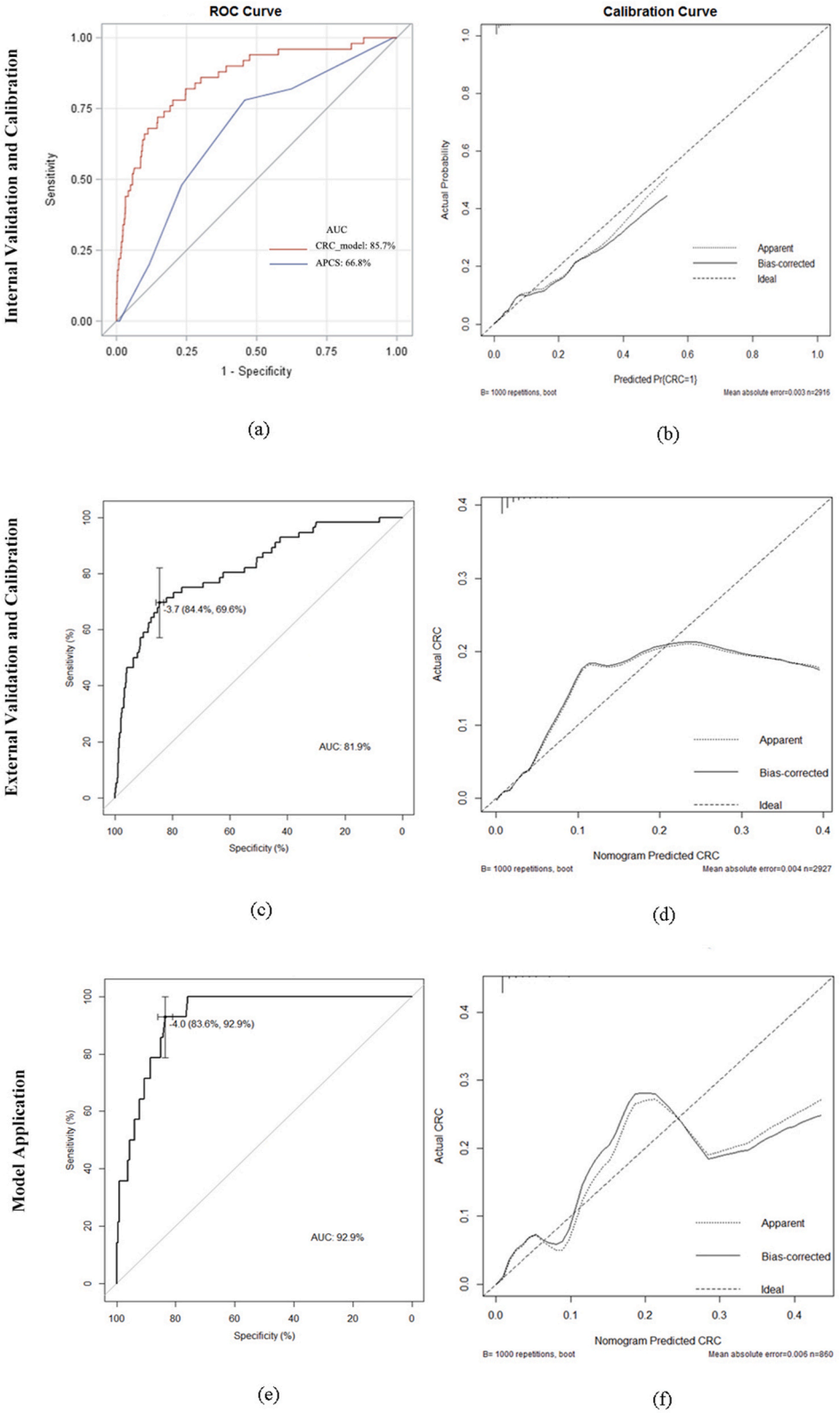

Fig. 3. Receiver operating curve (ROC) and a calibration curve of the nomogram-predicted CRC model.

Table 3
Area under the receiver operating characteristic curve (AU-ROC) (95 % confidence interval, CI) for predicting CRC.

	AUC	95%CI		*P*
		Lower	Upper	
APCS	0.6682	0.5968	0.7396	ref
CRC	0.8574	0.8006	0.9142	<0.0001

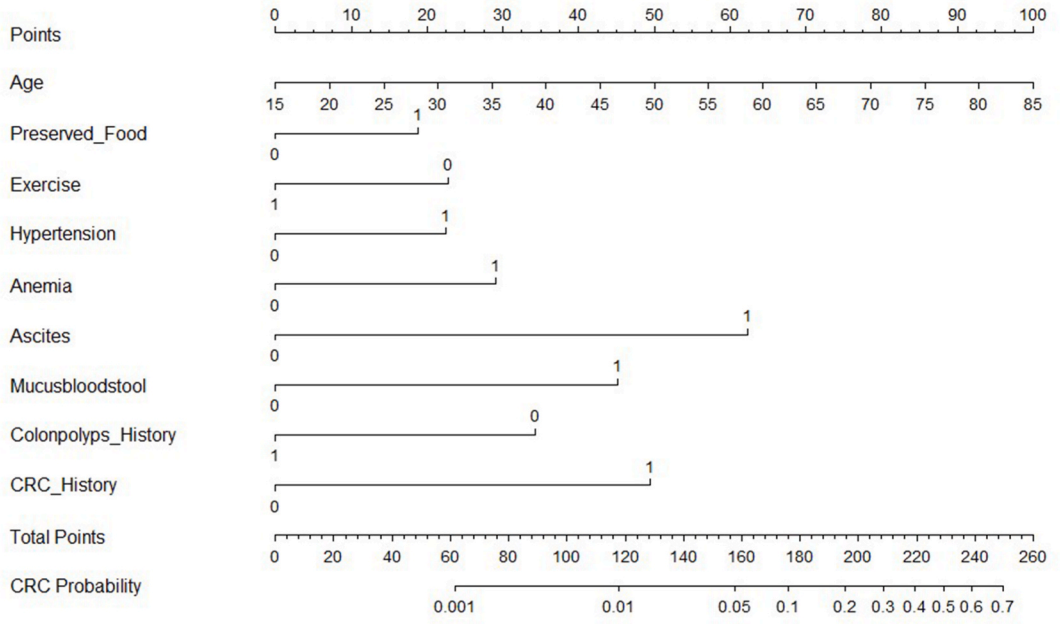

Fig. 4. Self-reported symptom-based nomogram for outpatient cases.

Table 4
The possibility of scoring by nomogram prediction and actual rate in symptomatic patients.

	Total	Non-CRC (model predicted)		CRC	
	N	N	%	N	%
<19	145	145	100	0	0
≤72	2286	2270	99.3	16	0.7
>72	630	596	94.6	34	5.4

3.4. Clinical application of the decision-making model

In Step 3, the "Clinical application of the decision-making model" includes two sequentially stages: user experience evaluation (n = 2690) and clinical practicality evaluation (n = 860). User experience was mainly assessed in terms of the out-patient's acceptance of the questionnaire and the investigator's proficiency in mastering the model scoring system. A total of 2390 valid questionnaires were collected for evaluating user experience from the outpatient departments of five hospitals across China from November 1, 2021, to

Table 5
Comparison of the consistency between model prediction and clinical decisions of outpatient doctors (undergo colonoscopy or not).

		Total	Consistent with Clinical Decisions	Inconsistent with Clinical Decisions
			(Proportion of Cohort, n %)	*(Proportion of Cohort, n %)*
Model prediction	High-risk	736 (85.6 %)	649(75.5 %)[a]	87(10.1 %)
	Low-risk	124 (14.4 %)	51(5.9 %)	73(8.5 %)
Total		860 (100 %)[b]	700(81.4 %)	160(18.6 %)

[a] All the 14 outpatients detected with CRC are included in this cell (14/649).
[b] The detection rate of CRC was 1.63 % (14/860).

January 6, 2022. A total of 99.9 % (2388 subjects) of the patients were satisfied with the questionnaire survey form, and none of the patients had difficulty understanding or answering. A total of 17 investigators participated in the model scoring, all of whom found the nomogram easy to grasp. Among 2390 patients, 860 outpatient cases received colonoscopy examination. The numbers are not equal because the number of patients who responded to the questionnaire is smaller than the number of patients who underwent colonoscopies, given the exclusion of inclusion/exclusion violations and consent decertification. Fourteen patients were confirmed to have CRC, and the detection rate of CRC was 1.63 %. A total of 700 out of 860 outpatient cases' model prediction results were consistent with the clinical decisions of outpatient doctors, and the total consistency ratio was 81.4 %. Application of this decision-making model detected 100 % of outpatient cases with CRC (n = 14), which were all in the high-risk group (nomogram score ≥72) defined by the model (Table 5).

There were 124 subjects in total in whom the decision-making model considered a low risk of CRC (nomogram score <72), even though 73 of their prediction results were inconsistent with the clinical decisions (doctors decided on colonoscopy for CRC screening). All patients underwent colonoscopy. Subsequent colonoscopy results confirmed that the model's predictions were reasonable and that none of the 124 outpatient cases were diagnosed with CRC. This finding proves that this predictive model could better distinguish low-risk outpatient cases, regardless of whether the clinical doctors judged them as low risk. In other words, 860 colonoscopies could be performed to detect 14 CRC outpatient cases in the 'colonoscopy for everyone' strategy, whereas the decision-making model strategy resulted in 736 colonoscopies to detect all of the CRC outpatient cases (Table 5). The marginal benefit of colonoscopy is minimal among persons categorized by the decision-making model as low risk.

The ROC (95 % CI) of this nomogram was 0.9290 (0.8892–0.9683), specificity = 92.9 % and sensitivity = 83.6 %. The calibration plot suggests a favorable agreement between the predictions and observations (MAE = 0.006). The comparison between Step 1 and Step 2 showed that the model's specificity, sensitivity, and accuracy were superior to those of the former 2 steps (Fig. 3).

4. Discussion

Considerable evidence has shown that the incidence and mortality of CRC could be reduced through early colonoscopy screening. However, the implementation of colonoscopy-based screening is usually confined by insufficient resources, low participant compliance, and concern about complication rates [15].

Upon evaluation of the current prediction model, we have identified two primary limitations. Firstly, it is not specifically designed for outpatient scenarios in which crucial information regarding symptoms and lifestyle risk factors is not adequately incorporated. Models like Imperiale's [10] and the Scottish Intercollegiate Guidelines Network (SIGN) referral criteria [12], despite providing clear guidelines for primary healthcare settings or cancer referrals, their assessment accuracy remains susceptible to individual variations among the target population and the specific application scenarios (outpatient). Secondly, the scoring systems of those models are complex and difficult to master, including the Harvard CRC Risk Assessment Tool [13] and the Freedman's Model [11], are overly intricate and challenging to master. Predominantly based on Western population data, they have limited applicability to diverse ethnic groups, and their intricate assessment processes necessitate the involvement of specialized medical personnel.

It is noteworthy that the APCS Score system [14], which shares similar design objectives with our model, is tailored to the cancer incidence patterns and population characteristics of the Asia-Pacific region, thus ensuring high regional applicability. However, due to its oversimplified design, which aims at enhancing practicality in the preclinical community screening stage, it fails to effectively capture the diverse risk factors for cancer, with only four factors included. Consequently, the assessment results often lack accuracy. Furthermore, the APCS Score system has not undergone external validation in the specific subgroup of outpatient patients, which limits its generalizability in this particular setting.

In contrast, our model innovatively adds clinical application and, after external verification, demonstrates simplicity and user-friendliness.

More importantly, successful translation of these findings into specific scenarios, especially assisting outpatient colonoscopy allocation, remains a major challenge.

In this study, the model was based entirely on patients' self-reported symptoms, and a nomogram was constructed that is easy to promote and apply. Fifty-six features were chosen based on the CRC consensus and literature concerning CRC was analyzed [4,21,22]. The AUC of CRC was significantly higher than that of the improved APCS (85.7 % vs 66.8 %), which indicated that the prediction ability of this model is better than that of APCS. This model covered exercise habits, preserved food intake, history of polyps, and important symptom information. The nomogram model's visual/quantitative features can cover important symptoms and lifestyle factors based on easy collection. To an extent, this model strikes a good balance between clinical utility and quantification.

The predictive ability of the model is important, while the practicality of the model and the compliance of patients are also important aspects worth evaluating. An accurate scoring system is somehow challenging to master because it pursues a higher sensitivity and specificity, leading to the collection of more information. The complexity of a scoring system might limit its use in clinical or community settings. Moreover, despite the excellent discrimination, it is inevitable to encounter factors that are difficult to measure or recall bias. Especially for new outpatient cases with symptoms, establishing a rapid auxiliary evaluation system is also more conducive to improving patient compliance. This decision-making model was developed based on symptoms reported by outpatient cases. Our data showed that patients could fill out the scoring form without difficulty and assess their risk level before choosing the colonoscopy. This model is conducive to helping doctors when deciding if colonoscopy is necessary. Participants with high-risk scores may be prioritized for screening colonoscopy. In contrast, participants with low-risk scores may still be encouraged to undergo less invasive screening tests than colonoscopy, such as fecal immunochemical test (FIT) [7,23].

Actually, this model was proven to better distinguish low-risk outpatients (none of the 124 low-risk patients were diagnosed with

CRC), regardless of whether the clinical doctors judged it or not, and decreased the inefficiency involved in looking for CRC. These risk-adapted screening strategies may improve the effectiveness and acceptance of current screening models. They reduce the burden of invasive procedures for low-risk populations while focusing on high-risk populations. Therefore, risk-adapted screening strategies may also improve the cost-effectiveness of current screening models. In fact, we are currently developing a mobile application that will allow patients to self-report their symptoms through a mobile application accessible via a QR code. Once completed, outpatient physicians will have access to the model's assessment results through their computer workstations. We believe that the integration of this application will significantly enhance the practicality and usability of our model.

There are limitations to this study. This model is designed for an optimal CRC screening strategy. Additionally, due to the relatively low incidence of CRC, prolonging the observation period of this prospective cohort would be more convincing. The incidence of colorectal diseases, including colorectal polyps, diverticular disease, melanosis coli, and inflammatory bowel disease, is increasing annually, according to an investigation of 350,000 cases [24–26]. These colorectal diseases are also considered indications for colonoscopy. Therefore, this model cannot replace the clinical interviews of specialist physicians. To improve the predictive power and clinical value of this model, further long-term prospective cohort studies will be needed.

5. Conclusion

This multicenter study developed and validated a simple and user-friendly decision-making CRC model that covered essential self-reported symptom information and lifestyle risk factors for patients. The CRC model was demonstrated to have good performance in outpatient decision-making and clinical utility, particularly because it could better differentiate CRC low-risk outpatient cases, regardless of whether the clinical doctors judged them as low risk.

Funding

This research was supported by the National Natural Science Foundation of China [32170898] and the Key project of Shanxi Science and Technology Department [201903D321030].

Study approval and ethics statement

This study was conducted ethically in accordance with the World Medical Association Declaration of Helsinki. The study protocol was approved by the Chinese PLA General Hospital Ethics Committee (approval number: 301hn11-2017-05). Written Informed consent was obtained from all patients for the publication of all their data and/or images.

Data availability statement

Data associated with this study will be made available on request.

CRediT authorship contribution statement

Zhe Luan: Writing – review & editing, Writing – original draft. **Fangfang Liu:** Formal analysis. **Li Zhang:** Conceptualization. **Jun Chen:** Data curation. **Yiming Zhao:** Investigation. **Congyong Li:** Investigation. **Zhaoyun Liu:** Investigation. **Huawei Li:** Methodology. **Li Dong:** Investigation. **Funing Zang:** Investigation. **Lingyan Han:** Software, Resources. **Tianyue Zhao:** Investigation. **Qiao Wang:** Investigation. **Gang Sun:** Conceptualization. **Shufang Wang:** Writing – review & editing, Validation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e33619.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (3) (2021) 209–249.

[2] M. Cao, H. Li, D. Sun, W. Chen, Cancer burden of major cancers in China: a need for sustainable actions, Cancer Commun. 40 (5) (2020) 205–210.

[3] W. Chen, R. Zheng, P.D. Baade, S. Zhang, H. Zeng, F. Bray, et al., Cancer statistics in China, 2015, CA A Cancer J. Clin. 66 (2) (2016) 115–132.

[4] R.A. Smith, K.S. Andrews, D. Brooks, S.A. Fedewa, D. Manassaram-Baptiste, D. Saslow, et al., Cancer screening in the United States, 2017: a review of current American Cancer Society guidelines and current issues in cancer screening, CA A Cancer J. Clin. 67 (2) (2017) 100–121.

[5] J.J. Yang, A. Thanataveerat, P.H. Green, B. Lebwohl, Cost effectiveness of routine duodenal biopsy analysis for celiac disease during endoscopy for gastroesophageal reflux, Clin. Gastroenterol. Hepatol. 13 (8) (2015) 1437–1443.

[6] S. Wang, Z. Wang, L. Peng, X. Zhang, J. Li, Y. Yang, et al., Gender, age, and concomitant diseases of melanosis coli in China: a multicenter study of 6,090 cases, PeerJ 6 (2018) e4483.

[7] L.M. Helsingen, P.O. Vandvik, H.C. Jodal, T. Agoritsas, L. Lytvyn, J.C. Anderson, et al., Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: a clinical practice guideline, Bmj 367 (2019) l5515.

[8] H.M. Chiu, J.Y. Ching, K.C. Wu, R. Rerknimitr, J. Li, D.C. Wu, et al., A risk-scoring system combined with a fecal immunochemical test is effective in screening high-risk subjects for early colonoscopy to detect advanced colorectal neoplasms, Gastroenterology 150 (3) (2016) 617–625.e613.

[9] G.A. Colditz, K.A. Atwood, K. Emmons, R.R. Monson, W.C. Willett, D. Trichopoulos, et al., Harvard report on cancer prevention volume 4: harvard cancer risk index. Risk index working group, harvard center for cancer prevention, Cancer Causes Control 11 (6) (2000) 477–488.

[10] T.F. Imperiale, D.R. Wagner, C.Y. Lin, G.N. Larkin, J.D. Rogge, D.F. Ransohoff, Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer, Ann. Intern. Med. 139 (12) (2003) 959–965.

[11] A.N. Freedman, M.L. Slattery, R. Ballard-Barbash, G. Willis, B.J. Cann, D. Pee, et al., Colorectal cancer risk prediction tool for white men and women without known susceptibility, J. Clin. Oncol. 27 (5) (2009) 686–693.

[12] J. Cubiella, M. Salve, M. Díaz-Ondina, P. Vega, M.T. Alves, F. Iglesias, et al., Diagnostic accuracy of the faecal immunochemical test for colorectal cancer in symptomatic patients: comparison with NICE and SIGN referral criteria, Colorectal Dis. 16 (8) (2014) O273–O282.

[13] B.A. Adelstein, P. Macaskill, R.M. Turner, P.H. Katelaris, L. Irwig, The value of age and medical history for predicting colorectal cancer and adenomas in people referred for colonoscopy, BMC Gastroenterol. 11 (2011) 97.

[14] M.C. Wong, T.Y. Lam, K.K. Tsoi, H.W. Hirai, V.C. Chan, J.Y. Ching, et al., A validated tool to predict colorectal neoplasia and inform screening choice for asymptomatic subjects, Gut 63 (7) (2014) 1130–1136.

[15] L. Peng, K. Weigl, D. Boakye, H. Brenner, Risk scores for predicting advanced colorectal neoplasia in the average-risk population: a systematic review and meta-analysis, Am. J. Gastroenterol. 113 (12) (2018) 1788–1800.

[16] R.L. Sutherland, D.J. Boyne, T.N. Jarada, L.M. Lix, J. Tinmouth, L. Rabeneck, et al., Development and validation of a risk prediction model for high-risk adenomas at the time of first screening colonoscopy among screening aged Canadians, Prev. Med. 148 (2021) 106563.

[17] F. Fernández-Bañares, R. Clèries, J. Boadas, J. Ribes, J.C. Oliva, A. Alsius, et al., Prediction of advanced colonic neoplasm in symptomatic patients: a scoring system to prioritize colonoscopy (COLONOFIT study), BMC Cancer 19 (1) (2019) 734.

[18] A.M. Carrington, D.G. Manuel, P. Fieguth, T.O. Ramsay, V. Osmani, B. Wernly, et al., Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation, IEEE Trans. Pattern Anal. Mach. Intell. (2022).

[19] K. Thanikachalam, G. Khan, Colorectal cancer and nutrition, Nutrients 11 (1) (2019).

[20] E.J. Kuipers, W.M. Grady, D. Lieberman, T. Seufferlein, J.J. Sung, P.G. Boelens, et al., Colorectal cancer, Nat. Rev. Dis. Prim. 1 (2015) 15065.

[21] H.W. Park, S. Han, J.S. Lee, H.S. Chang, D. Lee, J.W. Choe, et al., Risk stratification for advanced proximal colon neoplasm and individualized endoscopic screening for colorectal cancer by a risk-scoring model, Gastrointest. Endosc. 76 (4) (2012) 818–828.

[22] M. Betés, M.A. Muñoz-Navas, J.M. Duque, R. Angós, E. Macías, J.C. Súbtil, et al., Use of colonoscopy as a primary screening test for colorectal cancer in average risk people, Am. J. Gastroenterol. 98 (12) (2003) 2648–2654.

[23] H.M. Chiu, G.H. Jen, Y.W. Wang, J.C. Fann, C.Y. Hsu, Y.C. Jeng, et al., Long-term effectiveness of faecal immunochemical test screening for proximal and distal colorectal cancers, Gut 70 (12) (2021) 2321–2329.

[24] C.B. Aarons, S. Shanmugan, J.I. Bleier, Management of malignant colon polyps: current status and controversies, World J. Gastroenterol. 20 (43) (2014) 16178–16183.

[25] M.E.S. Bronzwaer, L. Koens, W.A. Bemelman, E. Dekker, P. Fockens, Volume of surgery for benign colorectal polyps in the last 11 years, Gastrointest. Endosc. 87 (2) (2018) 552–561.e551.

[26] G. Longcroft-Wheaton, P. Bhandari, Management of early colonic neoplasia: where are we now and where are we heading? Expet Rev. Gastroenterol. Hepatol. 11 (3) (2017) 227–236.