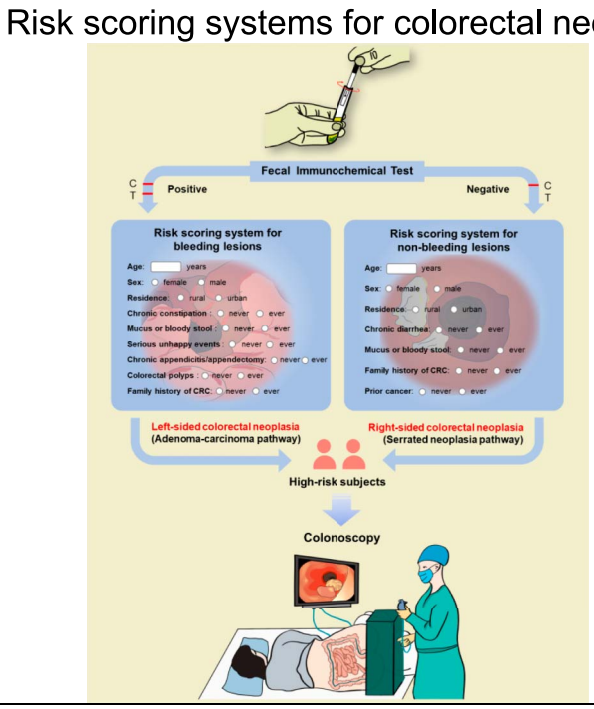
Risk Scoring Systems for Predicting the Presence of Colorectal Neoplasia by Fecal Immunochemical Test Results in Chinese Population

Weimiao Wu, PhD[1,*], Xin Chen, BA[1,*], Chen Fu, MD[2], Martin C.S. Wong, MD[3], Pingping Bao, PhD[2], Junjie Huang, PhD[3], Yangming Gong, MD[2], Wanghong Xu, PhD[1] and Kai Gu, MD[2]

INTRODUCTION: **Adherence to colonoscopy screening for colorectal cancer (CRC) is low in general populations, including those tested positive in the fecal immunochemical test (FIT). Developing tailored risk scoring systems by FIT results may allow for more accurate identification of individuals for colonoscopy.**

METHODS: **Among 807,109 participants who completed the primary tests in the first-round Shanghai CRC screening program, 71,023 attended recommended colonoscopy. Predictors for colorectal neoplasia were used to develop respective scoring systems for FIT-positive or FIT-negative populations using logistic regression and artificial neural network methods.**



Wu et al. *Clin Trans Gastroenterol.* [Month Year]. [doi:10.14309/ctg.0000000000000525]

Clinical and Translational GASTROENTEROLOGY

[1]Global Health Institute, School of Public Health, Fudan University, Shanghai, China; [2]Shanghai Municipal Center for Disease Control & Prevention, Shanghai, China; [3]The Jockey Club School of Public Health and Primary Care, Faculty of Medicine, Chinese University of Hong Kong, Hong Kong SAR, China.
Correspondence: Wanghong Xu, PhD. E-mail: wanghong.xu@fudan.edu.cn. Kai Gu, MD. E-mail: gukai@scdc.sh.cn.
*Weimiao Wu and Xin Chen contributed equally to this work.
Received June 21, 2022; accepted August 10, 2022; published online August 25, 2022

RESULTS: **Age, sex, area of residence, history of mucus or bloody stool, and CRC in first-degree relatives were identified as predictors for CRC in FIT-positive subjects, while a history of chronic diarrhea and prior cancer were additionally included for FIT-negative subjects. With an area under the receiver operating characteristic curve of more than 0.800 in predicting CRC, the logistic regression-based systems outperformed the artificial neural network-based ones and had a sensitivity of 68.9%, a specificity of 82.6%, and a detection rate of 0.24% by identifying 17.6% subjects at high risk. We also reported an area under the receiver operating characteristic curve of about 0.660 for the systems predicting CRC and adenoma, with a sensitivity of 57.8%, a specificity of 64.6%, and a detection rate of 6.87% through classifying 38.1% subjects as high-risk individuals. The performance of the scoring systems for CRC was superior to the currently used method in Mainland, China, and comparable with the scoring systems incorporating the FIT results.**

DISCUSSION: **The tailored risk scoring systems may better identify high-risk individuals of colorectal neoplasia and facilitate colonoscopy follow-up. External validation is warranted for widespread use of the scoring systems.**

SUPPLEMENTARY MATERIAL accompanies this paper at http://links.lww.com/CTG/A867

Clinical and Translational Gastroenterology 2022;13:e00525. https://doi.org/10.14309/ctg.0000000000000525

INTRODUCTION

Population-based colorectal cancer (CRC) screening has been applied in national healthcare service programs in various countries. The fecal immunochemical test (FIT), due to its affordability, convenience, and accuracy, is the most widely used primary screening test to identify individuals for further colonoscopy examination (1, 2). Despite the well-recognized effectiveness of the FIT colonoscopy cascade screening on the reduced CRC incidence and mortality worldwide (3–5), the use of FIT as a primary screening test may have missed quite a lot of cases with nonbleeding colorectal lesions. FIT is sensitive to bleeding lesions, but not to those nonbleeding. As colonoscopy is not recommended currently for those with negative FIT, more than 15% of CRC and 70% of precursor lesions could be missed by FIT screening alone (6). In a meta-analysis, the sensitivity of 1 or more FITs was observed around 75% for CRC but only 25% for advanced adenoma using a positivity threshold of 20 μg Hb/g feces (7). Regarding the subjects positive in FIT in observational studies, only 76.0% (range: 12.5%–96.7%) attended the recommended colonoscopy (8–10). In our previous report, we found that only 39.8% of high-risk individuals attended a colonoscopy follow-up in a Chinese population of Shanghai (11) and the adherence was as low as 24.0% in the population of Pudong New Area of Shanghai, in which only 30.4% of the FIT-positive subjects attended colonoscopy (12). To develop tailored risk assessment tools by FIT results may help to address the issues: for subjects with negative results, those with nonbleeding lesions may be identified as high-risk individuals for subsequent colonoscopy, and for those with positive results, further risk assessment may motivate them to attend colonoscopy and thus improve the adherence rate.

Nonbleeding colorectal lesions may differ in biological characteristics, progression, and carcinogenesis from those bleeding (13,14). It has been suggested that screen-detected CRC among FIT-positive people was usually left-sided and followed the adenoma-carcinoma pathway, while the interval CRC among those negative in FIT was more likely right-sided and to be serrated neoplasia (15–17). There exist different risk factors for bleeding and nonbleeding colorectal neoplasia. A family history of CRC was associated with advanced colorectal neoplasia among FIT-negative subjects (18,19), but not identified as a risk factor among FIT-positive populations (20,21). These findings justify the need for tailored risk scoring systems according to FIT results.

In this study, we aimed to develop and validate FIT-specific scoring systems for CRC in all screened subjects and for colorectal neoplasia in high-risk subjects attending colonoscopy. We also compared the performance of the FIT-specific scoring systems for CRC with (i) the parallel use of risk stratification and FIT, the recommended primary screening test for CRC in Mainland of China (22), and (ii) the unified scoring system incorporating FIT results in our previous report (23), both of which followed a traditional primary screening of risk assessment and FIT, ignoring the heterogeneity in biological characteristics, risk factors, and progression between the bleeding and nonbleeding cancers. Our analyses may help to develop a novel FIT-risk assessment primary screening, which may facilitate more accurate risk stratification for colonoscopy and improve the effectiveness and cost-effectiveness of CRC screening.

METHODS

We reported this cross-sectional study following the guidelines of the Strengthening the Reporting of Observational Studies in Epidemiology (24) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (25). The Shanghai CRC screening program was approved by the Ethics Review Committee of the Shanghai Municipal Center for Disease Control & Prevention. Written informed consent was obtained from all study participants.

Study population

A community-based CRC screening program was performed using the recommended screening strategy in the 16 districts of Shanghai, China, in 2013. As described in our previous reports, all permanent residents aged 50–74 years with no history of CRC were eligible and recruited to participate in the program voluntarily (11,23).

All participants were provided with a questionnaire-based risk stratification and 2-sample qualitative FIT for primary screening, followed by recommended colonoscopy for those positive in at least 1 assessment, similar to the screening strategy used in Jiashan County, Zhejiang Province of China (22). A unique 12-digital barcode was assigned to each participant to label the questionnaire, fecal samples collected, and colonoscopy examination result.

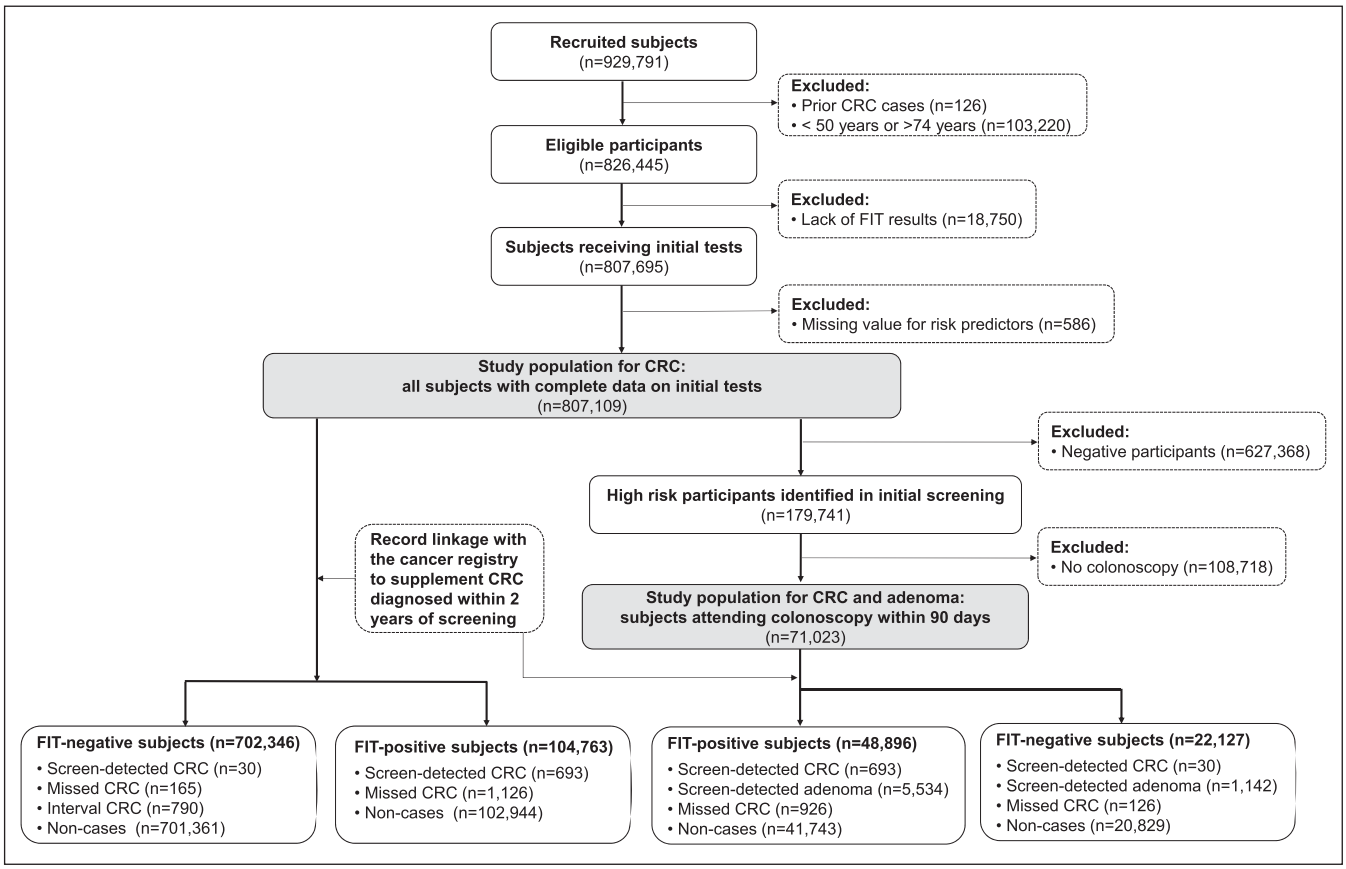


Figure 1. Flow diagram for recruitment of the study participants. CRC, colorectal cancer; FIT, fecal immunochemical test.

Data collection

Questionnaire-based risk stratification was conducted through face-to-face interviews. All participants were asked to provide sociodemographic information including date of birth, sex, educational level, marital status, occupation, area of residence, and to answer the questions for risk stratification. As previously reported (22), the subjects were regarded as high-risk for CRC if they had 1 of the 3 events: (i) diagnosis of any cancer, (ii) history of colorectal polyps, and (iii) CRC in first-degree relatives and/or had at least 2 of the 6 conditions: (i) history of chronic constipation, (ii) history of chronic diarrhea, (iii) history of mucus or bloody stool, (iv) serious unhappy life events that caused psychiatric trauma (e.g., divorce and death of any family member), (v) chronic appendicitis or appendectomy, and (vi) chronic cholecystitis or cholecystectomy. These risk factors were determined based on several epidemiological studies conducted in China between 1970s and 1990s (26) and proved to be effective and cost-effective for CRC screening in Chinese populations (11,22). The information on smoking, drinking, physical activity, dietary habits, and body mass index, the risk factors of CRC in other populations, was not collected in this study.

At the interview, all subjects were instructed to collect 2 fecal samples with an interval of 1 week. The samples were collected in tubes containing about 5 mL moist feces and were required to return to a local hospital within 48 hours. The FIT results were read within 5 minutes of colloidal gold assay, with a positivity threshold of 100 ng Hb/mL (20 μg Hb/g feces).

Participants who were positive in risk stratification and/or any FIT were referred to designated hospitals for colonoscopy examination of the entire colorectum. All lesions detected were further evaluated by biopsy to confirm the pathological diagnosis. When using CRC as the outcome among the 807,109 subjects with complete data of primary tests, we defined the prevalent CRC at the screening as those diagnosed within 2 years of screening to allow the clinical identification, as previously described (23). Therefore, the outcome in the analysis included the screen-detected CRC obtained from the program reporting system (within 90 days of screening), and the missed or interval CRC supplemented through a record linkage with the Shanghai Cancer Registry using a unique ID number (within 2 years of screening). For the outcome of CRC and adenoma, we restricted the study population to the 71,023 high-risk subjects who attended recommended colonoscopy. The screen-detected CRC and adenoma (within 90 days of screening) were obtained from the program reporting system for those attending colonoscopy in the designated hospitals of the program and from the Shanghai medical record system for those receiving colonoscopy outside. Similarly, the missed CRC (within 2 years of screening) was supplemented through a record linkage with the Shanghai Cancer Registry (Figure 1).

The Shanghai Municipal Center for Disease Control and Prevention was responsible for the management, supervision, and evaluation of the CRC screening program. The program staff attended the annual training course before screening. Specialists in surgery, endoscopy, and pathology were organized for quality assurance of the diagnosis and treatment of colorectal lesions.

Table 1. Demographic characteristics and risk factors of study participants by the results of 2-sample FIT

	All subjects (n = 807,109)			Subjects attending colonoscopy (n = 71,023)		
	Positive in either FIT (n = 104,763)	Negative in both FIT (n = 702,346)	*P* values[a]	Positive in either FIT (n = 48,896)	Negative in both FIT (n = 22,127)	*P* values[a]
Age at screening (yr), mean (SD)	62.4 (6.0)	61.8 (6.0)	<0.001	62.1 (5.9)	61.9 (5.9)	<0.001
Sex, n (%)			<0.001			<0.001
Men	45,025 (43.0)	271,043 (38.6)		20,911 (42.8)	7,778 (35.2)	
Women	59,738 (57.0)	431,303 (61.4)		27,985 (57.2)	14,349 (64.9)	
Educational level, n (%)			<0.001			<0.001
No formal education	10,762 (10.3)	52,734 (7.5)		5,287 (10.8)	1,982 (9.0)	
Primary school	35,418 (33.8)	195,543 (27.8)		17,404 (35.6)	5,398 (24.4)	
Middle school	50,666 (48.4)	387,442 (55.2)		22,801 (46.6)	12,188 (55.1)	
High school	7,734 (7.4)	65,211 (9.3)		3,337 (6.8)	2,521 (11.4)	
College or above	183 (0.2)	1,416 (0.2)		67 (0.1)	38 (0.2)	
Marital status, n (%)			<0.001			<0.001
Married	95,959 (91.6)	637,594 (90.8)		44,956 (91.9)	19,827 (89.6)	
Unmarried	2,128 (2.0)	19,374 (2.8)		972 (2.0)	537 (2.4)	
Divorced	979 (0.9)	7,718 (1.1)		410 (0.8)	335 (1.5)	
Widowed	4,920 (4.7)	29,772 (4.2)		2,220 (4.5)	1,237 (5.6)	
Unknown	777 (0.7)	7,888 (1.1)		338 (0.7)	191 (0.9)	
Occupation, n (%)			<0.001			<0.001
Office workers	6,344 (6.1)	47,351 (6.7)		2,772 (5.7)	1,890 (8.5)	
Enterprise workers	36,004 (34.4)	289,925 (41.3)		15,829 (32.4)	9,569 (43.3)	
Farmers	38,542 (36.8)	205,846 (29.3)		19,145 (39.2)	5,418 (24.5)	
Self-employed	2,664 (2.5)	16,830 (2.4)		1,391 (2.8)	574 (2.6)	
Unemployed	5,346 (5.1)	28,041 (4.0)		2,621 (5.4)	1,235 (5.6)	
Others	15,863 (15.1)	114,353 (16.3)		7,138 (14.6)	3,441 (15.6)	
Area of residence, n (%)			<0.001			<0.001
Urban	22,852 (21.8)	262,502 (37.4)		8,687 (17.8)	8,151 (36.8)	
Rural	81,911 (78.2)	439,844 (62.6)		40,209 (82.2)	13,976 (63.2)	
Factors for risk stratification, n (%)						
Chronic diarrhea	7,137 (6.8)	36,283 (5.2)	<0.001	3,776 (7.7)	5,463 (24.7)	<0.001
Chronic constipation	8,858 (8.5)	46,284 (6.6)	<0.001	4,361 (8.9)	5,514 (24.9)	<0.001
Mucus or bloody stool	3,462 (3.3)	13,882 (2.0)	<0.001	1,813 (3.7)	2,728 (12.3)	<0.001
Chronic appendicitis/appendectomy	12,771 (12.2)	66,705 (9.5)	<0.001	6,188 (12.7)	7,067 (31.9)	<0.001
Chronic cholecystitis/cholecystectomy	11,321 (10.8)	63,330 (9.0)	<0.001	5,616 (11.5)	7,502 (33.9)	<0.001
Serious unhappy life events	2,757 (2.6)	15,287 (2.2)	<0.001	1,320 (2.7)	2,307 (10.4)	<0.001
Diagnosis of any cancer	2,354 (2.3)	13,963 (2.0)	<0.001	1,109 (2.3)	3,470 (15.7)	<0.001
CRC in first-degree relatives	3,949 (3.8)	21,380 (3.0)	<0.001	2,146 (4.4)	6,235 (28.2)	<0.001
Previous colorectal polyps	2,055 (2.0)	9,457 (1.4)	<0.001	1,302 (2.7)	3,957 (17.9)	<0.001

CRC, colorectal cancer; FIT, fecal immunochemical test.
[a]*P* values for *t* test or χ^2 tests.

Statistical analysis

The subjects were randomized into a derivation set and a validation set at a ratio of 8:2 (27), with the number of subjects in the derivation versus the validation set being 83,812:20,951 in FIT-positive subjects and 561,871:140,475 in FIT-negative subjects to predict the outcome of CRC, and 39,117:9,779 and 17,702:4,425, respectively, to predict the outcome of CRC and adenoma.

Univariable and multivariable logistic regression (LR) analyses were applied to identify possible risk factors for colorectal neoplasia among FIT-positive and FIT-negative subjects, respectively.

Table 2. Univariable and multivariable LR analysis of risk factors for colorectal neoplasia, stratified by the results of 2-sample FIT

Risk factors	Subjects positive in either FIT				Subjects negative in both FIT			
	β (SE) in univariable analysis	*P* values	β (SE) in multivariable analysis	*P* values	β (SE) in univariable analysis	*P* values	β (SE) in multivariable analysis	*P* values
For CRC in all subjects								
Age at screening (yr)	0.058 (0.004)	<0.001	0.052 (0.004)	<0.001	0.059 (0.005)	<0.001	0.055 (0.005)	<0.001
Sex (men vs women)	0.516 (0.048)	<0.001	0.433 (0.048)	<0.001	0.394 (0.064)	<0.001	0.307 (0.064)	<0.001
Area of residence (urban vs rural)	0.494 (0.051)	<0.001	0.421 (0.052)	<0.001	0.381 (0.064)	<0.001	0.344 (0.064)	<0.001
Chronic diarrhea	0.290 (0.084)	<0.001	—	—	0.614 (0.111)	<0.001	0.526 (0.113)	<0.001
Chronic constipation	−0.020 (0.086)	0.813	—	—	0.184 (0.119)	0.121	—	—
Mucus or bloody stool	0.833 (0.093)	<0.001	0.705 (0.094)	<0.001	0.662 (0.168)	<0.001	0.378 (0.172)	0.028
Chronic appendicitis/ appendectomy	−0.020 (0.073)	0.788	—	—	−0.030 (0.110)	0.784	—	—
Chronic cholecystitis/ cholecystectomy	−0.027 (0.077)	0.731	—	—	0.197 (0.103)	0.056	—	—
Serious unhappy events	0.183 (0.136)	0.178	—	—	0.312 (0.189)	0.098	—	—
Diagnosis of any cancer	0.146 (0.149)	0.328	—	—	0.972 (0.145)	<0.001	0.814 (0.146)	<0.001
CRC in first-degree relatives	0.457 (0.102)	<0.001	0.373 (0.103)	<0.001	0.746 (0.132)	<0.001	0.659 (0.134)	<0.001
Colorectal polyps	0.065 (0.166)	0.693	—	—	0.728 (0.195)	<0.001	—	—
For CRC and adenoma in subjects attending colonoscopy								
Age at screening (yr)	0.039 (0.002)	<0.001	0.032 (0.002)	<0.001	0.037 (0.005)	<0.001	0.035 (0.005)	<0.001
Sex (men vs women)	0.674 (0.026)	<0.001	0.619 (0.026)	<0.001	0.554 (0.058)	<0.001	0.495 (0.058)	<0.001
Area of residence (urban vs rural)	0.213 (0.032)	<0.001	0.199 (0.033)	<0.001	−0.225 (0.061)	<0.001	−0.259 (0.062)	<0.001
Chronic diarrhea	−0.059 (0.049)	0.224	—	—	0.007 (0.066)	0.919	—	—
Chronic constipation	−0.309 (0.049)	<0.001	−0.267 (0.050)	<0.001	−0.156 (0.069)	0.023	—	—
Mucus or bloody stool	0.148 (0.065)	0.022	—	—	−0.078 (0.089)	0.383	—	—
Chronic appendicitis/ appendectomy	−0.068 (0.039)	0.082	−0.089 (0.040)	0.025	−0.059 (0.062)	0.340	—	—
Chronic cholecystitis/ cholecystectomy	−0.072 (0.041)	0.080	—	—	−0.100 (0.061)	0.102	—	—
Serious unhappy events	0.214 (0.074)	0.004	0.281 (0.075)	<0.001	0.074 (0.091)	0.417	—	—
Diagnosis of any cancer	−0.078 (0.088)	0.379	—	—	−0.086 (0.081)	0.286	—	—
CRC in first-degree relatives	0.167 (0.059)	0.005	0.169 (0.061)	0.005	0.177 (0.062)	0.004	0.245 (0.063)	<0.001
Colorectal polyps	−0.199 (0.085)	0.019	−0.300 (0.086)	<0.001	−0.086 (0.077)	0.259	—	—

CRC, colorectal cancer; FIT, fecal immunochemical test; LR, logistic regression.

Variables with *P* values less than 0.10 in the univariable analysis were then included in multivariable LR models with a backward elimination. Those remaining significant were used to develop the final models. The regression coefficients of selected variables were transferred into point values, with each point equivalent to the increase in the risk of colorectal neoplasia associated with a 5-year increase in age (i.e., the coefficient of age multiplied by 5) (28). The risk score for each subject was the sum of the point values of all included variables in the final model.

The ANN models using the multilayer perceptron method were also constructed based on the significant variables in the final multivariable LR models. The ANN architecture consists of an input layer (predictor variables), an output layer (outcome variable), and a hidden layer (latent variables connecting predictors and the outcome) (29). The relationship between input and output neurons was processed and weighted through the hidden layer. The weights derived from the ANN models were transferred into risk scores using formula (equation [1]),

Table 3. Scoring algorithm to calculate point values for LR-based FIT-specific scoring systems of colorectal neoplasia, by the results of 2-sample FIT

	For CRC in all subjects					For CRC and adenoma in subjects attending colonoscopy				
		Positive in either FIT		Negative in both FIT		Positive in either FIT			Negative in both FIT	
Risk factors	Reference value (W_{ij})	β (95% CI)[a]	Score[b]	β (95% CI)[a]	Score[b]	Reference value (W_{ij})	β (95% CI)[a]	Score[b]	Reference value (W_{ij})	β (95% CI)[a]	Score[b]
Age at screening (yr)		0.056 (0.047 to 0.065)		0.058 (0.046 to 0.069)			0.032 (0.027 to 0.037)			0.036 (0.025 to 0.047)	
Age group											
50–54	52 (W_{ref})	—	0	—	0	52 (W_{ref})	—	0	52 (W_{ref})	—	0
55–59	57	—	1.0	—	1.0	57	—	1.0	57	—	1.0
60–64	62	—	2.0	—	2.0	62	—	2.0	62	—	2.0
65–69	67	—	3.0	—	3.0	67	—	3.0	67	—	3.0
70–74	72	—	4.0	—	4.0	72	—	4.0	72	—	4.0
Sex											
Women	0 (W_{ref})	—	0	—	0	0 (W_{ref})	—	0	0 (W_{ref})	—	0
Men	1	0.416 (0.311 to 0.522)	1.0	0.315 (0.180 to 0.449)	1.0	1	0.614 (0.557 to 0.672)	4.0	1	0.503 (0.376 to 0.631)	3.0
Area of residence											
Rural	0 (W_{ref})	—	0	—	0	0 (W_{ref})	—	0	1	0.228 (0.093 to 0.363)	1.0
Urban	1	0.389 (0.275 to 0.503)	1.0	0.314 (0.180 to 0.449)	1.0	1	0.181 (0.109 to 0.253)	1.0	0 (W_{ref})	—	0
Chronic diarrhea											
Never	0 (W_{ref})	—	—	—	0	—	—	—	—	—	—
Ever	1	—	—	0.528 (0.292 to 0.764)	2.0	—	—	—	—	—	—
Mucus or bloody stool											
Never	0 (W_{ref})	—	0	—	0	—	—	—	—	—	—
Ever	1	0.610 (0.395 to 0.824)	2.0	0.333 (−0.034 to 0.700)	1.0	—	—	—	—	—	—
Chronic constipation											
Never	—	—	—	—	—	1	0.332 (0.220 to 0.445)	2.0	—	—	—
Ever	—	—	—	—	—	0 (W_{ref})	—	0	—	—	—
Chronic appendicitis/ appendectomy											
Never	—	—	—	—	—	1	0.089 (0.002 to 0.176)	1.0	—	—	—
Ever	—	—	—	—	—	0 (W_{ref})	—	0	—	—	—
Serious unhappy events											
Never	—	—	—	—	—	0 (W_{ref})	—	0	—	—	—
Ever	—	—	—	—	—	1	0.254 (0.089 to 0.419)	2.0	—	—	—

Table 3. (*continued*)

Risk factors	For CRC in all subjects					For CRC and adenoma in subjects attending colonoscopy					
		Positive in either FIT		Negative in both FIT		Positive in either FIT			Negative in both FIT		
	Reference value (W_{ij})	β (95% CI)[a]	Score[b]	β (95% CI)[a]	Score[b]	Reference value (W_{ij})	β (95% CI)[a]	Score[b]	Reference value (W_{ij})	β (95% CI)[a]	Score[b]
CRC in first-degree relatives											
Never	0 (W_{ref})	—	0	—	0	0 (W_{ref})		0	0 (W_{ref})	—	0
Ever	1	0.426 (0.204 to 0.647)	2.0	0.676 (0.398 to 0.955)	2.0	1	0.160 (0.027 to 0.293)	1.0	1	0.233 (0.096 to 0.371)	1.0
Colorectal polyps											
Never	—	—	—	—	—	1	0.264 (0.076 to 0.451)	2.0	—	—	—
Ever	—	—	—	—	—	0 (W_{ref})		0	—	—	—
Diagnosis of any cancer											
Never	0 (W_{ref})	—	—	—	0	—	—		—	—	—
Ever	1	—	—	0.740 (0.424 to 1.056)	3.0	—	—		—	—	—
Total	—	—	0 to 10.0	—	0 to 14.0	—	—	0 to 17.0	—	—	0 to 9.0

CI, confidence interval; CRC, colorectal cancer; FIT, fecal immunochemical test; LR, logistic regression.

[a]β (95% CI) derived from the multivariable LR model.

[b]LR-based risk score = $\beta \times (W_{ij} - W_{ref})/B$, where constant B is the number of regression unit equivalent to 1 point in the final risk score and was calculated by multiplying the β for age by 5, e.g., for the LR-based scoring system for CRC, in FIT-positive subjects, constant B = 0.056 × 5 = 0.280. Based on an age-standardized method, the point values of other variables were obtained with their corresponding regression coefficients dividing by 0.280 and rounding to the nearest whole number: 0.416/0.280 = 1.0 for sex, 0.389/0.280 = 1.0 for the area of residence, 0.610/0.280 = 2.0 for mucus or bloody stool, and 0.426/0.280 = 2.0 for CRC in first-degree relatives.

Table 4. Discrimination and calibration of the FIT-specific risk scoring systems in predicting colorectal neoplasia

Scoring systems	For CRC in all subjects				For CRC and adenoma in subjects attending colonoscopy			
	Derivation set (n = 645,683)		Validation set (n = 161,426)		Derivation set (n = 56,819)		Validation set (n = 14,204)	
	AUC (95% CI)	*P* values[a]	AUC (95% CI)	*P* values[a]	AUC (95% CI)	*P* values[a]	AUC (95% CI)	*P* values[a]
LR-based								
FIT-positive	0.630 (0.616–0.644)	0.899	0.626 (0.597–0.655)	0.970	0.614 (0.607–0.622)	0.007	0.613 (0.598–0.629)	0.436
FIT-negative	0.637 (0.619–0.655)	0.599	0.635 (0.582–0.687)	0.708	0.605 (0.587–0.622)	0.680	0.603 (0.570–0.636)	0.049
Overall	0.804 (0.794–0.814)	0.516	0.855 (0.836–0.874)	0.777	0.660 (0.654–0.667)	0.008	0.659 (0.645–0.672)	0.044
ANN-based								
FIT-positive	0.626 (0.612–0.640)	0.057	0.626 (0.597–0.656)	0.911	0.609 (0.602–0.617)	0.163	0.612 (0.596–0.627)	0.816
FIT-negative	0.632 (0.614–0.651)	0.003	0.632 (0.579–0.684)	<0.001	0.602 (0.584–0.619)	0.883	0.604 (0.571–0.638)	0.499
Overall	0.802 (0.793–0.812)	0.002	0.854 (0.835–0.874)	<0.001	0.657 (0.650–0.664)	0.264	0.658 (0.645–0.671)	0.820

ANN, artificial neural network; AUC, area under the receiver operating characteristic curve; CI, confidence interval; CRC, colorectal cancer; FIT, fecal immunochemical test; LR, logistic regression.
[a]*P* values for calibration based on the Hosmer-Lemeshow goodness-of-fit tests.

where n and m represent the numbers of predictors and hidden units, respectively, u_{ij} is defined as the absolute value of weight between predictor i and hidden unit j divided by the total absolute values of weights pointing to hidden unit j, v_j is defined as the absolute value of weight between hidden unit j and the outcome divided by the total absolute values of weights pointing to the outcome, C_i represents the contribution of predictor i on the outcome, and the sum of contributions of all predictors is 100% (30). Then, the scores of predictors were computed by multiplying the total score in the corresponding LR-based scoring system by the contributions of predictors in the ANN model.

$$C_i = \sum_{j=1}^{m} v_j u_{ij}, \quad i = 1, \ldots, n \tag{1}$$

The discriminatory ability for each developed scoring system was measured using the area under the receiver operating characteristic curve (AUC) and compared using the DeLong test, while the ability in calibration (consistency of predicted risk with observed risk) was tested by the Hosmer-Lemeshow goodness-of-fit test. Net reclassification improvement and integrated discrimination improvement were further calculated to compare the predictive values of LR-based and ANN-based scoring systems. The optimal cutoff values for the scoring systems were identified based on the Youden index, sensitivity, specificity, proportion of high-risk individuals, detection rate, number of individuals needed to screen, and number of colonoscopies needed to identify one case. The performance of scoring systems in the derivation set were further evaluated in the split-sample validation set.

Considering that most FIT-based screening programs performed one sample test, we conducted a sensitivity analysis by redefining FIT positive as positive in the first FIT. As a result, the number of FIT-positive and FIT-negative subjects was 68,354 and 738,755, respectively, for the outcome of CRC, and was 32,233 and 38,790, respectively, for CRC and adenoma. We also conducted sensitivity analyses by supplementing missed or interval CRC with those diagnosed within 1 year of screening derived from the Shanghai Cancer Registry and by using CRC and adenoma detected within 90 days as the outcome.

All statistical analyses were performed in SAS 9.4 (SAS Institute, Cary, NC) and R version 4.0.2. Two-sided *P* values of less than 0.05 were considered statistically significant.

RESULTS

A total of 807,109 participants provided complete information for risk stratification, including 104,763 FIT-positive subjects and 702,346 FIT-negative subjects who were used to develop risk predictive models and scoring systems for CRC. Among 179,741 high-risk individuals who were positive in risk stratification or any FIT and thus recommended for colonoscopy, only 71,023 (39.5%) attended the examination within 90 days, in which risk predictive models and scoring systems for CRC and adenoma were established in 48,896 FIT-positive and 22,127 FIT-negative subjects, respectively. A total of 723 subjects were screening-detected with CRC and 6,676 with colorectal adenoma, as presented in Figure 1. Additional 2,081 missed or interval CRCs (74.2%) reported in cancer surveillance within 2 years of screening were also included in the analysis for the outcome of CRC.

Characteristics of study participants by FIT results

As presented in Table 1, significant differences in demographic and risk factors were observed between FIT-positive and FIT-negative groups in all participants and among those attending colonoscopy (all $P < 0.001$). For all subjects, those with positive FIT were older, less educated, and more likely to be men, lived in rural area, and had risk factors for risk stratification. Similar results were observed for subjects attending colonoscopy on demographic factors, but much higher prevalence of all factors for risk stratification was observed in the FIT-negative group in which all subjects were positive in risk stratification.

Selected risk predictors for colorectal neoplasia

In our previous report, age at screening, sex, chronic diarrhea, mucus or bloody stool, CRC in first-degree relatives, and

Table 5. Performance of several primary screening tests for the detection of colorectal neoplasia among all eligible subjects

Primary screening test[a]	High-risk subjects, n (%)	Detection rate, n (%)	Sensitivity (95% CI), %	Specificity (95% CI), %	No. of subjects needed to screen for 1 case	No. of colonoscopies needed to detect 1 case
For CRC (n = 807,109)						
Parallel use of risk assessment and FIT	179,740 (22.3)	2,013 (0.25)	71.8 (70.1–73.5)	77.9 (77.8–78.0)	401	89
LR-based scoring system incorporating FIT results	109,583 (13.6)	1,844 (0.23)	65.8 (64.0–67.5)	86.6 (86.5–86.7)	438	59
LR-based FIT-specific scoring system	141,692 (17.6)	1,933 (0.24)	68.9 (67.2–70.6)	82.6 (82.5–82.7)	418	73
ANN-based FIT-specific scoring system	161,709 (20.0)	1,975 (0.24)	70.4 (68.7–72.1)	80.1 (80.1–80.2)	409	82
For CRC and adenoma (n = 71,023)[b]						
LR-based FIT-specific scoring system	27,031 (38.1)	4,881 (6.87)	57.8 (56.7–58.8)	64.6 (64.2–65.0)	15	6
ANN-based FIT-specific scoring system	28,071 (39.5)	4,917 (6.92)	58.2 (57.1–59.2)	63.0 (62.6–63.4)	14	6

ANN, artificial neural network; CI, confidence interval; CRC, colorectal cancer; FIT, fecal immunochemical test; LR, logistic regression.
[a]FIT, positive referring to positive in either FIT.
[b]Among subjects attending colonoscopy.

diagnosis of any cancer were identified as predictors for CRC (23). In stratified analysis by FIT results, urban residency was additionally identified as a predictor for CRC in both FIT-positive and FIT-negative groups, whereas a history of chronic diarrhea and diagnosis of any cancer were no longer associated with CRC in the FIT-positive group (Table 2). Among subjects attending colonoscopy, 8 variables (age, sex, area of residence, constipation, chronic appendicitis or appendectomy, serious unhappy life events, CRC in first-degree relatives, and colorectal polyps) were identified as predictors for colorectal neoplasia (CRC and adenoma) in FIT-positive subjects, while 4 factors (age, sex, area of residence, and CRC in first-degree relatives) were selected for FIT-negative subjects. It is of note that opposite associations were observed for the area of residence in FIT-positive and FIT-negative subjects, and a negative association was observed for chronic constipation (β: −0.267, $P < 0.001$), chronic appendicitis or appendectomy (β: −0.089, $P = 0.025$), and colorectal polyps (β: −0.300, $P < 0.001$) among subjects positive in FIT (Table 2).

Development and validation of predictive models and scoring systems

Based on the selected predictors, we developed predictive models and corresponding scoring systems using LR and ANN methods. As presented in Table 3, the regression coefficients of risk predictors for CRC ranged from 0.610 (0.395–0.824) to 0.056 (0.047–0.065) in FIT-positive subjects and from 0.740 (0.424–1.056) to 0.058 (0.046–0.069) in FIT-negative subjects, based on which FIT-specific risk scoring systems with scores of 0–10 and scores of 0–14 were developed, respectively. For CRC and adenoma, the highest regression coefficient was observed for sex, while the lowest for age at screening, yielding a scoring system of 0–17 for FIT-positive participants and a system of 0–9 scores for FIT-negative participants. The weights of the predictors in ANN models are presented in the Supplementary Digital Content (see Supplementary Table 1, http://links.lww.com/CTG/A867). The corresponding weights for CRC predictors ranged from 35.1% to 12.4% in FIT-positive subjects and from 20.2% to 9.2% in FIT-negative subjects, while those for predictors of colorectal neoplasia varied from 24.4% to 4.3% and from 37.5% to 11.9% in FIT-positive and FIT-negative subjects, respectively.

The developed FIT-specific risk scoring systems demonstrated good discrimination and calibration in predicting colorectal neoplasia. As presented in Table 4, the AUCs (95% confidence interval [CI]) of the LR-based FIT-specific systems in predicting CRC were 0.804 (0.794–0.814) (P for calibration = 0.516) in the derivation set and 0.855 (0.836–0.874) (P for calibration = 0.777) in the validation set, while those for CRC and adenoma were 0.660 (0.654–0.667) (P for calibration = 0.008) and 0.659 (0.645–0.672) (P for calibration = 0.044), respectively. The AUCs for ANN-based scoring systems were comparable with those of LR-based systems for CRC (P for heterogeneity = 0.166) but were slightly inferior for CRC and adenoma (P for heterogeneity <0.05). Regardless of the FIT results, however, the values of the LR-based systems in predicting CRC and colorectal neoplasia were better than the ANN-based systems, with all net reclassification improvement >0 and integrated discrimination improvement >0 (all P values < 0.01).

Performance of scoring systems for colorectal neoplasia

Presented in Table 5 are the performance of the parallel use of risk stratification and FIT, the unified LR-based scoring system incorporating FIT results, and the FIT-specific systems established in this study. Using the score "1" as the cutoff point for the LR-based FIT-specific system to predict CRC in FIT-positive subjects

and the score "6" in FIT-negative subjects, we observed a higher specificity, a smaller number of high-risk individuals for colonoscopy, and less colonoscopies needed to identify 1 CRC than the parallel use of risk stratification and FIT, but with a reduced sensitivity and a larger number of subjects needed to screen. The performance of the LR-based FIT-specific systems was slightly better than the ANN-based ones but was inferior to the unified LR-based system incorporating FIT results.

For CRC and adenoma, the LR-based FIT-specific scoring systems using the score of "9" for FIT-positive subjects and "7" for FIT-negative subjects as an optimal cutoff point achieved a sensitivity of 57.8% (95% CI: 56.7–58.8), a specificity of 64.6% (95% CI: 64.2–65.0), a detection rate of 6.87% by identifying 38.1% subjects as at high risk for colonoscopy, with 15 subjects needed to screen and 6 colonoscopies needed to detect 1 case. The performance of ANN-based systems was comparable, with a slightly higher sensitivity but a slightly lower specificity.

Sensitivity analysis

When redefining FIT-positive as positive in the first FIT, the selected predictors for CRC were similar to those in the main analysis. For CRC and adenoma, however, mucus or bloody stool was selected instead of colorectal polyps and family history of CRC in FIT-positive subjects, while a negative association was observed for almost all risk factors in FIT-negative subjects. Presented in the Supplementary Digital Content (see Supplementary Table 2 and 3, http://links.lww.com/CTG/A867) are the scoring systems developed by integrating these selected predictors for colorectal neoplasia.

In sensitivity analyses by supplementing missed or interval CRC with those diagnosed within 1 year of screening, or by using detected CRC and adenoma within 90 days of screening as the outcome, the selected predictors were similar to those in the main analysis (data not shown in the tables). Consistent with the main results, all the sensitivity analyses demonstrated better discrimination and calibration of LR-based FIT-specific scoring systems than the ANN-based ones (see Supplementary Table 4, http://links.lww.com/CTG/A867) and showed better performance of the LR-based FIT-specific systems in predicting colorectal neoplasia (see Supplementary Table 5, http://links.lww.com/CTG/A867). The performance of LR-based FIT-specific system was also found comparable with the unified LR-based system incorporating FIT results in predicting CRC.

DISCUSSION

In this study, we found that the FIT-specific scoring systems outperformed the currently used primary screening method in identifying high-risk individuals. More importantly, the novel FIT-risk assessment primary screening, proposed in this study enables more accurate identification of high-risk individuals and more efficient allocation of the limited colonoscopy and thereby bears a potential to improve effectiveness and cost-effectiveness of CRC screening.

Almost all previous risk predictive models or risk scoring systems for CRC or advanced colorectal neoplasia were established for the whole screening populations. These models usually included demographics, medical history, lifestyle habits, specific symptoms or signs, or even biomarkers as risk predictors (31–33). Although the risk factors significantly differed between subjects positive in FIT and those negative (19,34,35), no predictive model was specifically developed by the test results. So far, only 2 models have been established for FIT-positive subjects by integrating FIT concentration and other factors to predict the outcome of CRC or advanced neoplasia (20,21).

In this study, for the first time, we developed FIT-specific scoring systems for CRC and colorectal neoplasia and provided tailored risk assessment tools to identify individuals with possible bleeding or nonbleeding colorectal lesions. With an AUC of more than 0.800 in identifying CRC, the FIT-specific scoring systems, either LR-based or ANN-based, outperformed those in previous studies demonstrating an AUC of 0.53–0.75 (20,33,36), including the parallel use of risk stratification and FIT in the Shanghai CRC screening program. The higher specificity, less demand for colonoscopy, and less colonoscopies needed to detect 1 CRC of the FIT-specific scoring systems suggest the value of our systems in reducing unnecessary colonoscopies (37), improving adherence to colonoscopy (8), and thus replacing the parallel use of risk stratification and FIT. Interestingly, we found that the performance of the FIT-specific scoring systems was comparable with the unified LR-based scoring system incorporating 2-sample FIT results (23), indicating that both scoring systems can be adopted in the population.

The LR-based FIT-specific scoring systems also performed well in predicting CRC and adenoma, with an AUC of 0.660. The AUC of our systems was within the AUCs of 0.62–0.75 for other scoring systems in predicting advanced colorectal neoplasia (20,21,31,38,39) and slightly higher than 0.62 of a scoring system for both advanced and nonadvanced neoplasia (40). Similar to previous studies (20,21), however, our systems for colorectal neoplasia were derived from high-risk individuals who attended colonoscopy and thus cannot be directly applied to general populations. Moreover, due to lack of detailed information to distinguish advanced adenoma from those nonadvanced, we used CRC and all adenomas as the outcomes, which is not consistent with those having using CRC and advanced adenomas as outcomes (31,33). Our systems may lead to a higher level of colonoscopy utilization in practice and compromise the cost-effectiveness of the screening program. However, it is estimated that close to 30% of those with nonadvanced adenoma (but no advanced neoplasia) at the age of 55 years would develop CRC in the absence of adenoma detection and removal (41). The information regarding the combined risk of CRC and adenoma may be of interest to those who might wish to have their nonadvanced adenomas removed (40). Moreover, colorectal serrated lesions, a kind of nonadvanced adenoma in adults, particularly among those negative in FIT, were found to progress more rapidly and lead to approximately 10%–20% of CRC cases (17). All of these support the benefits of detecting nonadvanced adenomas and indicate the desirability of extra colonoscopy examinations.

It is of note that chronic constipation, chronic appendicitis or appendectomy, and prior polyp seemed as protective factors for CRC and adenoma in FIT-positive population. The results were partly consistent with the report of Hreinsson et al. (42), in which CRC patients with overt or occult bleeding were less likely to have symptoms of constipation, diarrhea, and obstruction compared with nonbleeders. It is possible that subjects with constipation or appendicitis are at a higher risk of nonbleeding colorectal neoplasia, posing a competing risk for bleeding neoplasia (43). The instant polypectomy and regular medical examinations for

subjects with a history of colorectal polyp may help to explain the protective effect of this factor.

Evidently, the FIT-specific scoring systems developed in this study have important clinical and public health implications. First, the systems provide a simple-to-use risk assessment tool after a FIT, the most widely used screening test worldwide, facilitating risk assessment of colorectal neoplasia by screened subjects or healthcare providers. Second, the novel FIT-risk assessment primary screening, proposed in this study help to improve colonoscopy adherence in FIT-positive subjects through enhancing their awareness of CRC risk and allow for more accurate risk stratification in FIT-negative subjects. Finally, the scoring systems can be optimized by adjusting the cutoff points for consideration of cost-effectiveness, enabling policymakers to select the most suitable screening modality.

This study has several strengths. First, the large sample size provides enough statistical power to develop and validate the FIT-specific predictive models and scoring systems. Second, the data in the CRC screening program were collected under supervision, and the colonoscopy findings were available for all subjects in the analysis for CRC and adenoma. The high quality and the integrity of the data contribute to the accuracy of our results. Third, we reported this cross-sectional study following the guidelines of the Strengthening the Reporting of Observational Studies in Epidemiology and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis statements. Specifically, we established the scoring systems using an age-standardized method for LR coefficients and using weights from ANN models, which ensures the validity and interpretability of the scoring systems. Finally, the adjustable cutoff values of FIT-specific scoring systems may facilitate their utilities in screening practices according to the availability of healthcare resources.

There are several limitations that deserve attention. First, since the subjects in our study were screening volunteers, selection bias could not be excluded. However, this concern is partly mitigated by the relatively large sample size included in our analysis. Second, several important risk factors of colorectal neoplasia, such as smoking, body mass index, drinking, physical inactivity, and poor dietary habits included in previous studies (31,38), were not collected in our screened subjects. These unmeasured variables may decrease the predictive values of the developed risk scoring systems. However, smoking and body mass index, the 2 common risk factors for CRC, were not significantly associated with advanced colorectal neoplasia in the Chinese population (44), and sex may be a proxy for smoking due to the huge difference in smoking rate between Chinese men and women (45). Third, more than half of the high-risk subjects identified by the initial tests did not undergo scheduled colonoscopy in the program. The less representative study population may have underestimated the current risk of adenoma in the population and limited the generalization of our scoring systems for CRC and adenoma. For the outcome of CRC, the factors associated with the interval cancers may be different from those most strongly associated with the screen-detected CRC, which may have attenuated the accuracy of the scoring systems. However, in this study, we aimed to develop respective risk scoring systems by FIT results, which should be based on all the prevalent CRCs, including the interval cases. Finally, the predictive models and scoring systems were validated internally. External validation is warranted to extend the generalizability of our findings to other population groups.

In conclusion, the established LR-based scoring systems integrating specific risk factors for bleeding or nonbleeding colorectal neoplasia provide a more refined risk assessment in screened subjects having FIT. The novel FIT-risk assessment primary screening, may help to identify high-risk individuals for necessary colonoscopy and improve the efficacy of CRC screening. Further investigations are warranted to externally validate the scoring systems and evaluate the effectiveness and cost-effectiveness of the FIT-risk assessment primary screening, in large-scale population-based programs.

CONFLICTS OF INTEREST

Guarantor of the article: Wanghong Xu, MD, PhD.

Specific author contributions: Conception and design: W.X. and K.G.; collection and assembly of the data: C.F., P.B., Y.G.; analysis and interpretation of the data: W.W., X.C., M.C.S.W., J.H.; drafting of the manuscript: W.W. and X.C.; critical revision of the manuscript for important intellectual content: W.X., M.C.S.W., J.H.; final approval of the manuscript: All authors.

Financial support: This study was supported by the Health Commission of the Pudong New Area of Shanghai (No. PW2019A-5), the National Key R&D Program of China (No. 2017YFC1308800), and the Key Technology Research for Colorectal Cancer Screening and High-risk Population Follow-up (No. 20DZ1100103).

Potential competing interests: None to report.

Study Highlights

WHAT IS KNOWN

- ✓ Nonbleeding colorectal lesions differed in biological characteristics, progression, carcinogenesis, and risk factors from those bleeding.
- ✓ Fecal immunochemical test (FIT) is sensitive to bleeding lesions, but not to nonbleeding neoplasia.
- ✓ Suboptimal adherence to colonoscopy was observed among participants of colorectal cancer screening, including those with positive FIT.

WHAT IS NEW HERE

- ✓ This is the first study to develop tailored risk scoring systems among subjects positive or negative in FIT.
- ✓ The developed FIT-specific scoring systems outperformed the currently used initial screening method in Mainland, China.
- ✓ The novel FIT-risk assessment primary screening, provided more refined identification of individuals with bleeding or nonbleeding lesions.
- ✓ The novel FIT-risk assessment primary screening, can be applied to improve effectiveness of screening and colonoscopy adherence.
- ✓ The developed risk scoring systems can be used by screened subjects or healthcare providers.
- ✓ The cutoff points of the risk scoring systems can be adjusted for a cost-effectiveness consideration.

ACKNOWLEDGEMENT

We thank the healthcare staff and the participants of the first-round Shanghai Colorectal Cancer Screening Program.

REFERENCES

1. Navarro M, Nicolas A, Ferrandez A, et al. Colorectal cancer population screening programs worldwide in 2016: An update. World J Gastroenterol 2017;23:3632–42.
2. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: A global overview of existing programmes. Gut 2015;64:1637–49.
3. Lauby-Secretan B, Vilahur N, Bianchini F, et al, International Agency for Research on Cancer Handbook Working Group. The IARC perspective on colorectal cancer screening. N Engl J Med 2018;378:1734–40.
4. Giorgi Rossi P, Vicentini M, Sacchettini C, et al. Impact of screening program on incidence of colorectal cancer: A cohort study in Italy. Am J Gastroenterol 2015;110:1359–66.
5. Zorzi M, Fedeli U, Schievano E, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut 2015;64:784–90.
6. Lin JS, Perdue LA, Henrikson NB, et al. Screening for colorectal cancer: Updated evidence report and systematic review for the US preventive services task force. JAMA 2021;325:1978–98.
7. Imperiale TF, Gruber RN, Stump TE, et al. Performance characteristics of fecal immunochemical tests for colorectal cancer and advanced adenomatous polyps: A systematic review and meta-analysis. Ann Intern Med 2019;170:319–29.
8. Wu W, Huang J, Yang Y, et al. Adherence to colonoscopy in cascade screening of colorectal cancer: A systematic review and meta-analysis. J Gastroenterol Hepatol 2022;37:620–31.
9. Ruggli M, Stebler D, Gasteiger M, et al. Experience with a colorectal cancer campaign in Swiss pharmacies. Int J Clin Pharm 2019;41:1359–64.
10. Kapidzic A, Van Der Meulen MP, Van Roon AH, et al. Gender differences in fecal immunochemical test performance for early detection of colorectal neoplasia. Gastroenterology 2013;1:S599.
11. Gong Y, Peng P, Bao P, et al. The implementation and first-round results of a community-based colorectal cancer screening program in Shanghai, China. Oncologist 2018;23:928–35.
12. Wu WM, Wang Y, Jiang HR, et al. Colorectal cancer screening modalities in Chinese population: Practice and lessons in Pudong new area of Shanghai, China. Front Oncol 2019;9:399.
13. van der Meulen MP, Lansdorp-Vogelaar I, van Heijningen EMB, et al. Nonbleeding adenomas: Evidence of systematic false-negative fecal immunochemical test results and their implications for screening effectiveness-A modeling study. Cancer 2016;122:1680–8.
14. Loeve F, Boer R, van Oortmarssen GJ, et al. Impact of systematic false-negative test results on the performance of faecal occult blood screening. Eur J Cancer 2001;37:912–7.
15. Kang X, Zhang R, Kwong TN, et al. Serrated neoplasia in the colorectum: Gut microbiota and molecular pathways. Gut Microbes 2021;13:1–12.
16. Schreuders EH, Grobbee EJ, Nieuwenburg SAV, et al. Multiple rounds of one sample versus two sample faecal immunochemical test-based colorectal cancer screening: A population-based study. Lancet Gastroenterol Hepatol 2019;4:622–31.
17. Dekker E, Tanis PJ, Vleugels JLA, et al. Colorectal cancer. Lancet 2019; 394:1467–80.
18. de Klerk CM, Vendrig LM, Bossuyt PM, et al. Participant-Related risk factors for false-positive and false-negative fecal immunochemical tests in colorectal cancer screening: Systematic review and meta-analysis. Am J Gastroenterol 2018;113:1778–87.
19. Cheng WC, Chen PJ, Kang JW, et al. Age, male sex, smoking and metabolic syndrome as risk factors of advanced colorectal neoplasia for fecal immunochemical test negative patients. J Formos Med Assoc 2022; 121:402–8.
20. Thomsen MK, Pedersen L, Erichsen R, et al. Risk-stratified selection to colonoscopy in FIT colorectal cancer screening: Development and temporal validation of a prediction model. Br J Cancer 2022;126: 1229–35.
21. Cooper JA, Parsons N, Stinton C, et al. Risk-adjusted colorectal cancer screening using the FIT and routine screening data: Development of a risk prediction model. Br J Cancer 2018;118:285–93.
22. Meng W, Cai SR, Zhou L, et al. Performance value of high risk factors in colorectal cancer screening in China. World J Gastroenterol 2009;15: 6111–6.
23. Wu WM, Gu K, Yang YH, et al. Improved risk scoring systems for colorectal cancer screening in Shanghai, China. Cancer Med 2022;11: 1972–83.
24. von Elm E, Altman DG, Egger M, et al, STROBE Initiative. The strengthening the reporting of observational studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. PLoS Med 2007;4:e296.
25. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual Prognosis or diagnosis (TRIPOD): The TRIPOD statement. Ann Intern Med 2015;162:55–63.
26. Wu W, Huang J, Tan S, et al. Screening methods for colorectal cancer in Chinese populations. Hong Kong Med J 2022;28:183–5.
27. Wang Z, Lin F, Ma H, et al. Contrast-enhanced spectral mammography-based radiomics nomogram for the prediction of neoadjuvant chemotherapy-insensitive breast cancers. Front Oncol 2021;11:605230.
28. Sullivan LM, Massaro JM, D'Agostino RB, et al. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med 2004;23:1631–60.
29. Renganathan V. Overview of artificial neural network models in the biomedical domain. Bratisl Lek Listy 2019;120:536–40.
30. Gao R, Zhang S, Liu B. Variable selection approach based on neural network analysis. J Syst Eng 1998:34–9.
31. Peng L, Weigl K, Boakye D, et al. Risk scores for predicting advanced colorectal neoplasia in the average-risk population: A systematic review and meta-analysis. Am J Gastroenterol 2018;113:1788–800.
32. Yu L, Zhao G, Wang L, et al. A systematic review of microbial markers for risk prediction of colorectal neoplasia. Br J Cancer 2022;126:1318–28.
33. Ma GK, Ladabaum U. Personalizing colorectal cancer screening: A systematic review of models to predict risk of colorectal neoplasia. Clin Gastroenterol Hepatol 2014;12:1624–34.e1.
34. Longstreth GF, Anderson DS, Zisook DS, et al. Low rate of cancer detection by colonoscopy in asymptomatic, average-risk subjects with negative results from fecal immunochemical tests. Clin Gastroenterol Hepatol 2020;18:2929–36.e1.
35. Colussi D, Fabbri M, Zagari RM, et al. Lifestyle factors and risk for colorectal polyps and cancer at index colonoscopy in a FIT-positive screening population. United Eur Gastroenterol J 2018;6:935–42.
36. Jeon J, Du M, Schoen RE, et al, Colorectal Transdisciplinary Study and Genetics and Epidemiology of Colorectal Cancer Consortium. Determining risk of colorectal cancer and starting age of screening based on lifestyle, environmental, and genetic factors. Gastroenterology 2018; 154:2152–64.e19.
37. Ladabaum U, Dominitz JA, Kahi C, et al. Strategies for colorectal cancer screening. Gastroenterology 2020;158:418–32.
38. Imperiale TF, Monahan PO, Stump TE, et al. Derivation and validation of a predictive model for advanced colorectal neoplasia in asymptomatic adults. Gut 2021;70:1155–61.
39. Tao S, Hoffmeister M, Brenner H. Development and validation of a scoring system to identify individuals at high risk for advanced colorectal neoplasms who should undergo colonoscopy screening. Clin Gastroenterol Hepatol 2014;12:478–85.
40. Wong MCS, Lam TYT, Tsoi KKF, et al. A validated tool to predict colorectal neoplasia and inform screening choice for asymptomatic subjects. Gut 2014;63:1130–6.
41. Brenner H, Altenhofen L, Stock C, et al. Natural history of colorectal adenomas: Birth cohort analysis among 3.6 million participants of screening colonoscopy. Cancer Epidemiol Biomarkers Prev 2013;22:1043–51.
42. Hreinsson JP, Jonasson JG, Bjornsson ES. Bleeding-related symptoms in colorectal cancer: A 4-year nationwide population-based study. Aliment Pharmacol Ther 2014;39:77–84.
43. Noordzij M, Leffondre K, van Stralen KJ, et al. When do we need competing risks methods for survival analysis in nephrology? Nephrol Dial Transpl 2013;28:2670–7.
44. Chen H, Li N, Ren J, et al, Group of Cancer Screening Program in Urban China CanSPUC. Participation and yield of a population-based colorectal cancer screening programme in China. Gut 2019;68:1450–7.
45. Chen Z, Peto R, Zhou M, et al, China Kadoorie Biobank CKB Collaborative Group. Contrasting male and female trends in tobacco-attributed mortality in China: Evidence from successive nationwide prospective cohort studies. Lancet 2015;386:1447–56.