Cancer Biol Med 2026. doi: 10.20892/j.issn.2095-3941.2025.0702

LETTER

Enhancing FIT-based colorectal cancer screening through risk prediction: a comparative evaluation of 17 models in a large Chinese cohort

Le Wang[1]*, Weimiao Wu[1]*, Bin Liu[1,2]*, Chen Zhu[1], Yumeng Ding[1,3], Yi Zhou[1,2], Weiwei Chen[3], Lijuan Dai[3], Xueni Chen[3], Yuefan Shen[3], Tingting Pan[1], Xiaohui Sun[2], Huizhang Li[1], Dong Hang[3], Yingying Mao[2], Partha Basu[4], Lingbin Du[1,5]

[1]Department of Cancer Prevention, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou 310022, China; [2]Department of Epidemiology, School of Public Health, Zhejiang Chinese Medical University, Hangzhou 310053, China; [3]Department of Epidemiology, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing 211166, China; [4]Early Detection, Prevention & Infections Branch, International Agency for Research on Cancer (WHO), Lyon 69007, France; [5]Zhejiang Key Laboratory of Prevention, Diagnosis and Therapy for Gastrointestinal Cancer, Hangzhou 310022, China

Colorectal cancer (CRC), the third most common cancer and the second leading cause of cancer-related death globally, is responsible for more than 1.9 million new cases and 0.9 million deaths reported annually[1]. Randomized controlled trials and cohort studies have demonstrated that screening can significantly decrease CRC incidence and mortality[2]. Current screening guidelines primarily recommend CRC screening for individuals at average risk, typically starting at predefined ages[2]. However, because of individual variations in background CRC risk, a uniform screening approach might have limited effectiveness at the population level. Tailored screening strategies aimed at providing screening recommendations based on individual CRC risk have therefore been proposed to decrease the load on colonoscopy services and improve the benefit to harm ratio[3,4]. Several CRC risk prediction models have been developed to identify individuals at high risk of colorectal neoplasia. However, substantial gaps remain in the evidence, including model heterogeneity, bias, and a lack of external validation[5-8]. Few studies have simultaneously validated multiple risk prediction models within the same screening population, yet such validation is crucial for evaluating their performance and optimizing screening strategies.

Overall study design and analytical framework

To address this gap, we conducted a comprehensive head-to-head comparison of 17 previously published CRC risk prediction models within the large, real-world Zhejiang Provincial CRC Screening Program, which has been described in prior studies (**Supplementary material**)[9,10]. The screening program was approved by the Ethics Committee and Institutional Review Board of Zhejiang Cancer Hospital (approval No. IRB2023-464), and all participants provided written informed consent to participate. In the present study, all analyses were restricted to individuals with positive fecal immunochemical test (FIT) results who subsequently underwent diagnostic colonoscopy. Risk prediction models were applied to further stratify risk within this population. The outcome was advanced colorectal neoplasia (ACN), defined as CRC, advanced adenomas, or advanced serrated lesions. The model performance metrics, including the area under the receiver operating characteristic curve (AUC) and net reclassification improvement, were estimated. Detailed descriptions of variable handling, statistical tests, and uncertainty estimation are provided in the **Supplementary material**.

Detailed inclusion and exclusion criteria are provided in **Figure 1A**, and the basic characteristics of the validation cohort are presented in **Table S1**. Among the participants,

*These authors contributed equally to this work.

Correspondence to: Yingying Mao, Partha Basu and Lingbin Du
E-mail: myy@zcmu.edu.cn, basup@iarc.who.int and dulb@zjcc.org.cn
ORCID ID: https://orcid.org/0000-0003-3644-9160, https://orcid.org/0000-0003-0124-4050 and https://orcid.org/0000-0002-1992-7784

Received November 28, 2025; accepted February 26, 2026; published online March 24, 2026.
Available at www.cancerbiomed.org

©2026 The Authors. Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

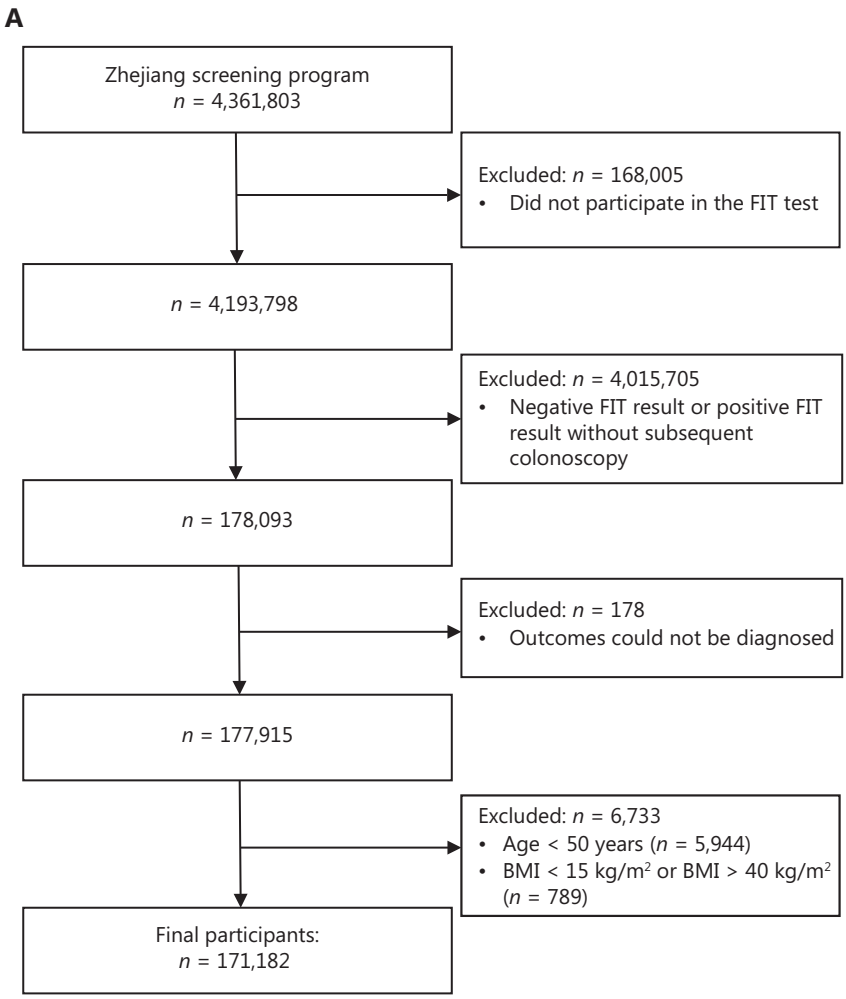


Figure 1 Continued

49.43% were men (n = 84,621), and approximately 2.42% reported a family history of CRC (n = 4,143). Except for vegetable and bean intake, which showed no significant differences ($P > 0.05$), all other factors demonstrated statistically significant differences between those with and without ACN ($P < 0.05$). A comparison of baseline characteristics between the present cohort and the general population revealed no significant differences, indicating that the cohort was representative of the broader population (**Tables S2 and S3**).

Identification and characteristics of colorectal cancer risk prediction models

Risk models were identified *via* synthesis of 4 previous systematic reviews[5-8] and a supplemental PubMed search covering January 2023 to October 2024 (**Figure 1B**). Briefly, we included models developed in screening populations that incorporated sociodemographic and environmental risk factors assessed by questionnaires, and predicted colorectal cancer or advanced adenomas confirmed by colonoscopy and pathology; detailed inclusion and exclusion criteria are provided in the **Supplementary material**. Basic characteristics and model information are presented in **Tables S4–S6**. Among the 17 models, approximately one-third were developed in Chinese populations, whereas the remainder originated predominantly from Europe and East Asia. Most models were developed in case–control settings, although 2 were derived from cohort studies. Across models, the predictor variables ranged from 3 to 9, and typically included age, sex, body mass index (BMI), smoking, alcohol consumption, and family history. The main characteristics of the validation cohort are shown in **Table S1**. The missing rates of questionnaire-based covariates were very low (0.0%–3.29%).

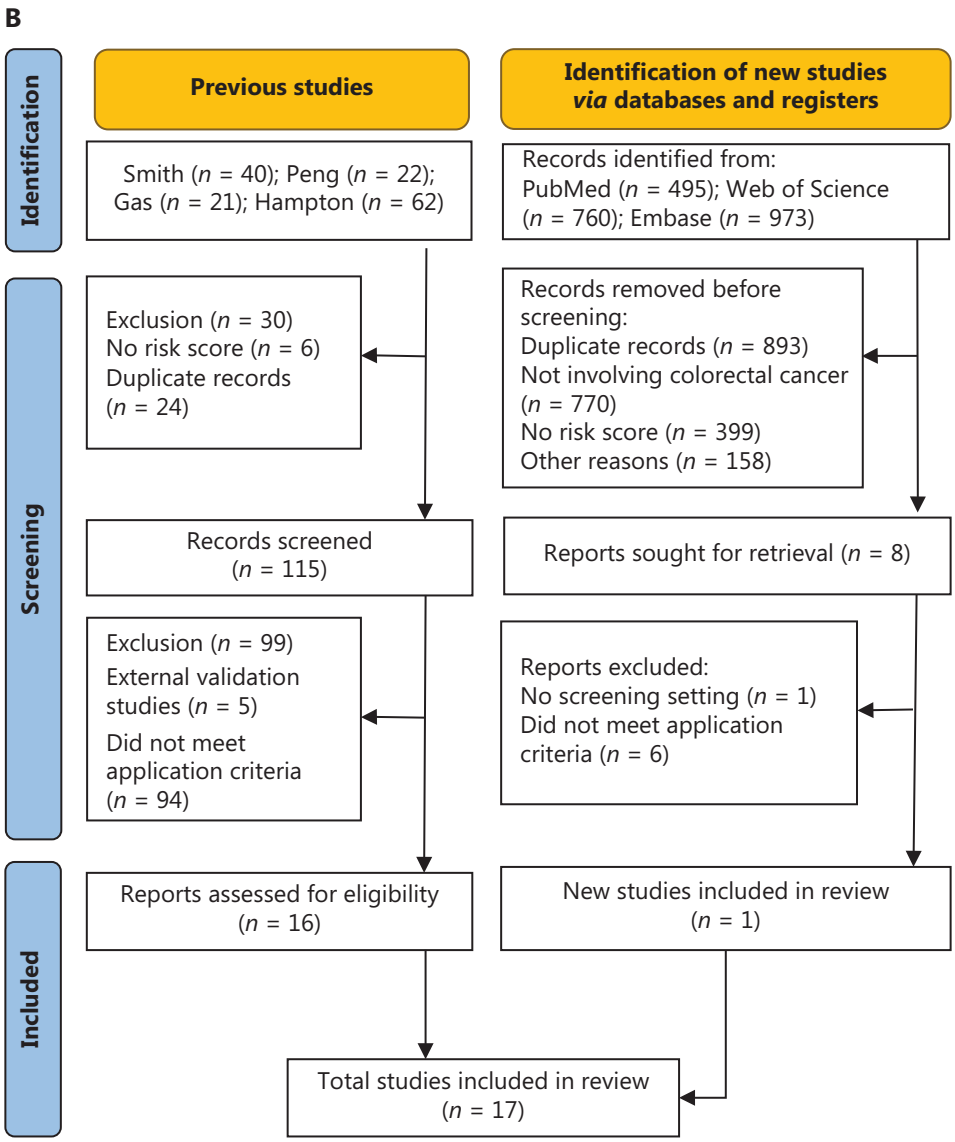


Figure 1 Continued

Therefore, missing values were not imputed; each covariate was treated as an exposure, and analyses were conducted on available data. Standardized mean difference analyses showed no statistically significant differences in demographic characteristics or lifestyle factors between FIT-positive individuals and the overall screening population (**Table S2**), or between colonoscopy attendees and non-attendees among the FIT-positive participants (**Table S3**), indicating that the validation cohort remained representative. Although development AUCs varied (0.60–0.74), validation outside the derivation populations was limited, heterogeneity in predictor definitions and model structure underscored the need for comparative validation using a standardized population and outcome definition.

Discrimination and calibration performance of risk prediction models in external validation

After exclusion of 2 models that did not report original AUC, the 15 models included in our study showed original AUCs ranging from 0.60 to 0.74 in the development, whereas the AUCs in the validation ranged from 0.60 to 0.65, indicating moderate discrimination ability between cases and non-cases. Notably, the AUCs were generally lower in the validation than the development (**Figure 1C, Table S7**). Across models, recalibration slopes generally <1 indicated attenuation of risk gradients in the external

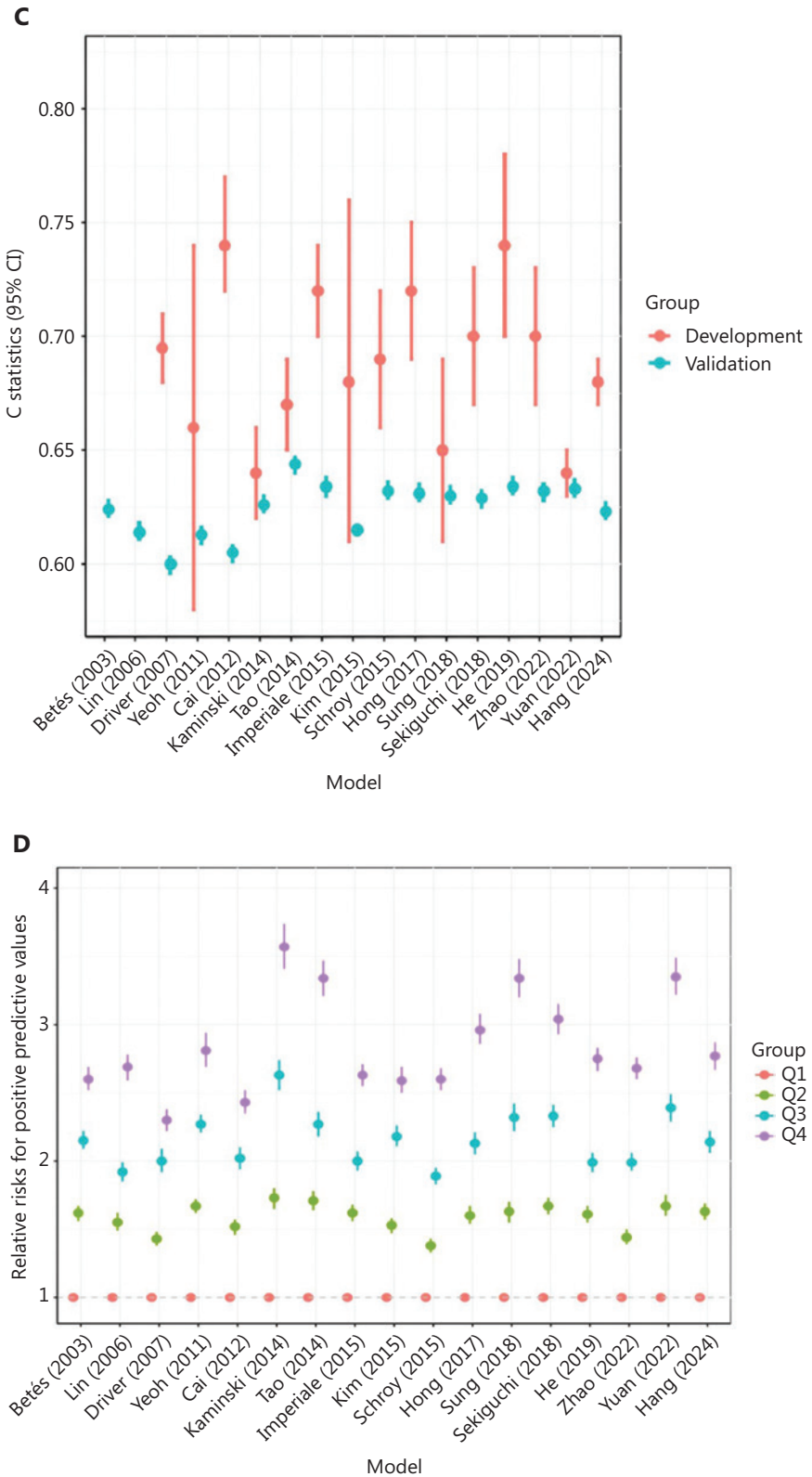


Figure 1 Continued

E

Model (year)	High risk (%)	No. of ACN (%)	PPV of ACN		
			Absolute rate, %	Absolute increase, %	Relative increase, %
Total	171,182 (100)	25,259 (100)	14.76	Reference	Reference
Betés (2003)	22,858 (13.35)	5,374 (21.28)	23.51	8.75	59.33
Lin (2006)	123,470 (72.13)	21,252 (84.14)	17.21	2.46	16.65
Driver (2007)	37,116 (21.68)	8,147 (32.25)	21.95	7.19	48.76
Yeoh (2011)	43,701 (25.53)	9,760 (38.64)	22.33	7.58	51.36
Cai (2012)	126,518 (73.91)	21,253 (84.14)	16.80	2.04	13.84
Kaminski (2014)	78,295 (45.74)	15,731 (62.28)	20.09	5.34	36.16
Tao (2014)	60,041 (35.07)	13,271 (52.54)	22.10	7.35	49.80
Imperiale (2015)	15,368 (8.98)	4,064 (16.09)	26.44	11.69	79.22
Kim (2015)	66,879 (39.07)	13,746 (54.42)	20.55	5.80	39.29
Schroy (2015)	73,240 (42.78)	15,130 (59.9)	20.66	5.90	40.00
Hong (2017)	146,498 (85.58)	23,560 (93.27)	16.08	1.33	8.99
Sung (2018)	89,618 (52.35)	17,479 (69.2)	19.50	4.75	32.18
Sekiguchi (2018)	52,152 (30.47)	11,572 (45.81)	22.19	7.43	50.38
He (2019)	106,271 (62.08)	19,644 (77.77)	18.48	3.73	25.27
Zhao (2022)	13,353 (7.8)	3,357 (13.29)	25.14	10.38	70.38
Yuan (2022)	37,644 (21.99)	8,946 (35.42)	23.76	9.01	61.06
Hang (2024)	88,958 (51.97)	16,825 (66.61)	18.91	4.16	28.18

Figure 1 (A) Flow diagram of the Zhejiang Colorectal Cancer Screening Program. (B) Flow diagram of systematic review. (C) Discrimination of the prediction models in development and external validation. Orange bars represent the C statistics of each model in the development phase, and blue bars represent the C statistics in the validation phase. Circles indicate point estimates, whereas horizontal lines denote 95% confidence intervals. (D) Relative risks of advanced colorectal neoplasia by quartile for each risk model. Circles indicate point estimates, whereas horizontal lines denote 95% confidence intervals. (E) Enrichment efficiency of risk-prediction models. FIT, fecal immunochemical test; BMI, body mass index; Q, quartile; ACN, advanced colorectal neoplasia; PPV, positive predictive value.

screening population, whereas recalibration intercepts reflected differences in baseline risk between the derivation and validation cohorts. This attenuation might have reflected differences in validation context, including the restriction of the study population to FIT-positive individuals and variations in population characteristics and outcome definitions with respect to the original model development settings. **Table S8** provides pairwise comparisons of net reclassification indices across the 17 models. Overall, most models showed no statistically significant differences in the net reclassification index, with values ranging from −0.16 to 0.29. Net reclassification improvement estimates were generally small, and many pairwise comparisons indicated confidence intervals including 0 (**Table S8**). These findings indicated limited evidence of meaningful improvement in individual risk reclassification. When stratified by lesion

subtype (**Table S7**), models originally developed to predict CRC demonstrated higher discrimination for CRC outcomes than advanced adenomas, whereas models designed for advanced adenoma showed more balanced performance across lesion types. These findings were consistent with the original design intent of the respective models. In sex-stratified analyses (**Table S9**), discrimination was consistently higher among women than men across most models, although the relative ranking of models was largely preserved between sexes. No individual model demonstrated clear superiority in one sex over the other.

Screening yield and risk stratification performance for ACN

Stratification of the risk scores showed a clear separation between the highest and lowest risk groups across models (**Figure 1D**). With respect to the lowest risk group, the positive predictive values (PPVs) of ACN were 11.36%–15.79% for the Q2 group, 16.11%–21.56% for the Q3 group, and 21.82%–26.68% for the Q4 group, corresponding to relative risks (RRs) of 1.38–1.73, 1.89–2.63, and 2.30–3.57, respectively. The PPVs and RRs for ACN and its subgroups by quartiles are presented in **Tables S10 and S11.**

Among all FIT-positive individuals undergoing colonoscopy, the overall PPV of ACN was 14.76%. When model-specific cutoffs were applied, the integration of risk stratification into FIT-based screening resulted in meaningful absolute increases of 1.33%–11.69% in the PPV of ACN, with relative improvements of 8.99%–79.22%. This approach also improved the enrichment efficiency by 9%–79%, indicating substantial practical gains in targeting high-risk individuals and optimizing colonoscopy allocation (**Figure 1E**). For example, with the risk model developed by Imperiale et al.[11], the PPV for ACN among FIT-positive individuals increased from 14.76% under the FIT-only strategy to 26.44% in the model-defined high-risk group, corresponding to an absolute increase of 11.69% and a relative improvement of 79.22%.

In conclusion, this study comprehensively evaluated the screening performance of various risk prediction models within a large-scale colorectal cancer screening program. The risk prediction models based on environmental risk factors demonstrated moderate effectiveness in identifying ACN. Incorporating these models into the FIT-positive population improved both the PPV and screening efficiency. The prominent roles of age and sex across models underscored their fundamental importance in colorectal cancer risk assessment. Therefore, a risk-stratified strategy, relative to the traditional age-based screening strategy, has potential to substantially enhance screening efficiency. The moderate discrimination observed might reflect the limited information captured by environmental risk factors alone, and the future integration of biomarkers or genomic variables might further improve risk stratification performance. Although existing risk prediction models may perform more discriminatively in women, their relative comparative performance remains broadly consistent between sexes. From a public health perspective, prioritizing risk models that yield favorable enrichment efficiency with a limited number of easily obtainable variables might offer promising pragmatic balance between effectiveness and scalability in population-based screening programs.

Grant support

This study was supported by Zhejiang Provincial Medical and Health Science and Technology Plan (Grant No. 2025KY677); Zhejiang Provincial Natural Science Foundation of China (Grant Nos. LTGY23H260004 and LQN25G030004); Noncommunicable Chronic Diseases-National Science and Technology Major Project (Grant Nos. 2024ZD0520100 and 2024ZD0520103); Postdoctoral Science Foundation of Zhejiang Province (Grant No. ZJ2024074); Science and Technology Program for Disease Prevention and Control of Zhejiang Province (Grant No. 2025JK018); and the UICC Technical Fellowship and Zhejiang Key Laboratory of Prevention; Diagnosis and Therapy for Gastrointestinal Cancer.

Conflict of interest statement

For authors identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article, which do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization.

Author contributions

Concept and design: Lingbin Du, Partha Basu, Yingying Mao, Le Wang, Weimiao Wu, Bin Liu.
Acquisition, analysis, or interpretation of data: Lingbin Du, Le Wang, Weimiao Wu, Chen Zhu, Tingting Pan, Huizhang Li.
Statistical analysis: Bin Liu, Yumeng Ding, Yi Zhou, Weiwei Chen, Lijuan Dai, Xueni Chen, Yuefan Shen.
Drafting of the manuscript: Lingbin Du, Partha Basu, Yingying Mao, Le Wang, Weimiao Wu, Bin Liu, Xiaohui Sun, Dong Hang.
Supervision: Lingbin Du, Partha Basu.
All authors confirm that they reviewed and critically revised the manuscript for important intellectual content, and approved the final draft for submission.

Data availability statement

The data generated in this study are available on request from the corresponding author.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024; 74: 229-63.
2. US Preventive Services Task Force; Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, Davis EM, et al. Screening for colorectal cancer: US Preventive Services Task Force recommendation statement. JAMA. 2021; 325: 1965-77.
3. Lansdorp-Vogelaar I, Meester R, de Jonge L, Buron A, Haug U, Senore C. Risk-stratified strategies in population screening for colorectal cancer. Int J Cancer. 2022; 150: 397-405.
4. Wang L, Liu C, Wang Y, Du L. Cost-effectiveness of risk-tailored screening strategy for colorectal cancer: a systematic review. J Gastroenterol Hepatol. 2022; 37: 1235-43.
5. Kang JHE, Levine E, Fleet A, Padilla MS, Lee JK, Harrison H, et al. Systematic review: risk prediction models for metachronous advanced colorectal neoplasia after polypectomy. J Gastroenterol Hepatol. 2024; 39: 2533-44.
6. Usher-Smith JA, Walter FM, Emery JD, Win AK, Griffin SJ. Risk prediction models for colorectal cancer: a systematic review. Cancer Prev Res (Phila). 2016; 9: 13-26.
7. Peng L, Weigl K, Boakye D, Brenner H. 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.
8. Hampton JS, Kenny RPW, Rees CJ, Hamilton W, Eastaugh C, Richmond C, et al. The performance of FIT-based and other risk prediction models for colorectal neoplasia in symptomatic patients: a systematic review. EClinicalMedicine. 2023; 64: 102204.
9. Hang D, Sun D, Du L, Huang J, Li J, Zhu C, et al. Development and evaluation of a risk prediction tool for risk-adapted screening of colorectal cancer in China. Cancer Lett. 2024; 597: 217057.
10. Hang D, Zhu C, Yang X, He J, Li H, Pan T, et al. Colorectal cancer screening based on fecal immunochemical test and risk assessment: a population-based study including two million participants in China. J Epidemiol. 2025; 35: 297-302.
11. Imperiale TF, Monahan PO, Stump TE, Glowinski EA, Ransohoff DF. Derivation and validation of a scoring system to stratify risk for advanced colorectal neoplasia in asymptomatic adults: a cross-sectional study. Ann Intern Med. 2015; 163: 339-46.

Cite this article as: Wang L, Wu W, Liu B, Zhu C, Ding Y, Zhou Y, et al. Enhancing FIT-based colorectal cancer screening through risk prediction: a comparative evaluation of 17 models in a large Chinese cohort. Cancer Biol Med. 2026; 23: 1031-1037. doi: 10.20892/j.issn.2095-3941.2025.0702