Chen *et al. Genome Medicine* (2026) 18:48
https://doi.org/10.1186/s13073-026-01623-z

Genome Medicine

RESEARCH Open Access

# Optimizing colorectal cancer screening through polygenic risk score-based risk stratification: evidence from a population-based cohort and screening trial

Hongda Chen[1†], Jiahui Si[2†], Chenyu Luo[1,3†], Jianbo Tian[4†], Junyi Xin[5†], Canqing Yu[6,7,8], Dianjianyi Sun[6,7,8], Pei Pei[8], Ling Yang[9,10], Iona Y. Millwood[9,10], Robin G. Walters[9,10], Yiping Chen[9,10], Huaidong Du[9,10], Zhengming Chen[10], Liming Li[6,7,8], Masaru Koido[11], Yoichiro Kamatani[11], Meilin Wang[12*], Xiaoping Miao[4,14,15*], Jun Lv[6,7,8,13*] and Min Dai[3*]

## Abstract

**Background** Accurate risk stratification of colorectal cancer (CRC) is essential for precise screening. Polygenic risk scores (PRS) hold promise for improving predictive efficacy in CRC. However, the real-world applicability of a risk-adapted CRC screening strategy based on the PRS remains underexplored. Therefore, we aimed to evaluate the optimized PRS in a large prospective cohort in China and assess its utility for risk-adapted CRC screening.

**Methods** We evaluated multiple PRS construction strategies using East Asian genome-wide association study data and well-established PRSs to select an optimal score, which was then assessed in 100,639 eligible participants from the China Kadoorie Biobank. The risk-adapted screening strategy assigns high-risk individuals to colonoscopy and low-risk individuals to fecal immunochemical testing (FIT), with FIT-positive cases referred for colonoscopy. We assessed the screening performance of the PRS, Asia–Pacific Colorectal Screening score, and their combination-based risk-adapted screening strategies against a standard FIT-based strategy among 2,821 participants in the TARGET-C CRC screening trial.

**Results** The combined PRS (i.e., $PRS_{121}$) demonstrated the best predictive performance (C-index = 0.602) for CRC. Individuals in the highest PRS quintile (top 20%) exhibited a 2.69-fold increased CRC risk compared with those in the lowest quintile. The high PRS and unfavorable lifestyle group conferred the highest risk (hazard ratio = 3.32, 95%

[†]Hongda Chen, Jiahui Si, Chenyu Luo, Jianbo Tian and Junyi Xin contributed equally to this work.

*Correspondence:
Meilin Wang
mwang@njmu.edu.cn
Xiaoping Miao
xpmiao@whu.edu.cn
Jun Lv
lvjun@bjmu.edu.cn
Min Dai
daimin2002@hotmail.com

Full list of author information is available at the end of the article

confidence interval: 1.80–6.11). In the TARGET-C trial, the $PRS_{121}$ effectively distinguished patients with advanced neoplasia from controls (top 20%, odds ratio = 3.66). The PRS-based risk-adapted screening strategy improved AN detection compared with FIT-only screening, with higher sensitivity (69.4% vs. 54.1%, $P = 7.6 \times 10^{-4}$) and detection rate (16.7% vs. 13.1%, $P = 0.024$). Notably, PRS-based screening identified 21.1% of AN cases that were missed by FIT, and integrated risk stratification using PRS and APCS increased this proportion to 37.9%, demonstrating substantial improvement in detecting FIT-negative lesions.

**Conclusions** PRS enables effective risk stratification for colorectal cancer and improves detection of advanced neoplasia by identifying high-risk individuals missed by FIT, supporting its utility in precision risk-adapted screening.

**Trial registration** Chinese Clinical Trial Registry: ChiCTR1800015506. Registered on 4 April 2018.

**Keywords** Colorectal cancer, Risk prediction, Polygenic risk score, Precise screening

## Background

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and the second leading cause of cancer-related mortality worldwide [1]. In countries where population-based CRC screening programs have been widely implemented, both the incidence and mortality of CRC have steadily declined [2, 3]. Current guidelines recommend endoscopic (e.g., colonoscopy) and stool-based tests (e.g., the fecal immunochemical test [FIT]) for CRC screening, typically for individuals above a certain age (usually 45 years) or at an elevated risk, such as those with a family history (FH) of CRC [4]. However, the one-size-fits-all screening strategy has yielded suboptimal performance, including low compliance rates, limited detection rates, and a high demand for endoscopic resources.

In the era of precision medicine, risk-adapted screening strategies have been proposed to identify individuals with a higher-risk profile who are more likely to harbor advanced colorectal neoplasia and benefit the most from targeted prevention [5]. We previously conducted a large-scale randomized controlled trial (RCT) in China that demonstrated that a risk-adapted screening approach, referring high-risk individuals for colonoscopy and low-risk individuals for FIT, achieved good compliance, reduced the colonoscopy demand, and maintained detection rates comparable to one-time colonoscopy over three screening rounds [6].

The effectiveness of risk-adapted strategies largely depends on the accuracy of the risk assessment tools. However, traditional tools rely primarily on demographic and lifestyle factors, demonstrating only a modest discriminative ability in detecting colorectal neoplasms. Over the past few decades, genome-wide association studies (GWAS) have identified hundreds of common variants associated with CRC risk [7, 8]. Polygenic risk scores (PRS) aggregate the effects of multiple genetic variants into a single quantifiable score and provide a promising approach for individualized CRC risk assessment [9]. However, few studies have systematically validated PRSs for CRC risk prediction in East Asian (EAS) populations, particularly in large-scale prospective cohorts [10, 11]. Moreover, the real-world applicability of the PRS, particularly its integration into risk-adapted screening strategies, remains insufficiently explored.

To bridge these gaps, we leveraged multiple independent EAS GWAS datasets to develop an optimized PRS and validated its performance using the China Kadoorie Biobank (CKB). Additionally, we established a comprehensive risk-adapted screening strategy incorporating colonoscopy and FIT and evaluated its effectiveness for detecting colorectal neoplasms in a real-world screening setting.

## Methods

### Study population

This study incorporated multiple independent datasets across different analytical stages, including PRS development, prospective assessment, and evaluation of screening performance. Specifically, we used (1) case–control genome-wide association study (GWAS) datasets for PRS construction, (2) a large prospective cohort for PRS assessment, and (3) a population-based CRC screening trial to evaluate the real-world performance of PRS-based risk-adapted screening strategies (Fig. 1).

#### *Case–control GWAS datasets for PRS development*

For PRS development, we used GWAS data from two independent CRC studies conducted in China together with publicly available GWAS summary statistics from Japan. The Chinese datasets included the Nanjing GWAS (1,316 CRC cases and 2,207 controls) and the Wuhan GWAS (3,000 CRC cases and 3,000 controls), which have been described previously [12, 13]. In brief, in Wuhan GWAS, CRC cases were recruited from Renmin Hospital of Wuhan University and diagnosed based on clinical and pathological records, while controls were selected from a contemporaneous community nutrition survey, matched to cases by sex and age (± 5 years). In Nanjing GWAS, the CRC cases were recruited from the Cancer Center of Nanjing Medical University, and the cancer-free controls were from the same districts of Nanjing. All

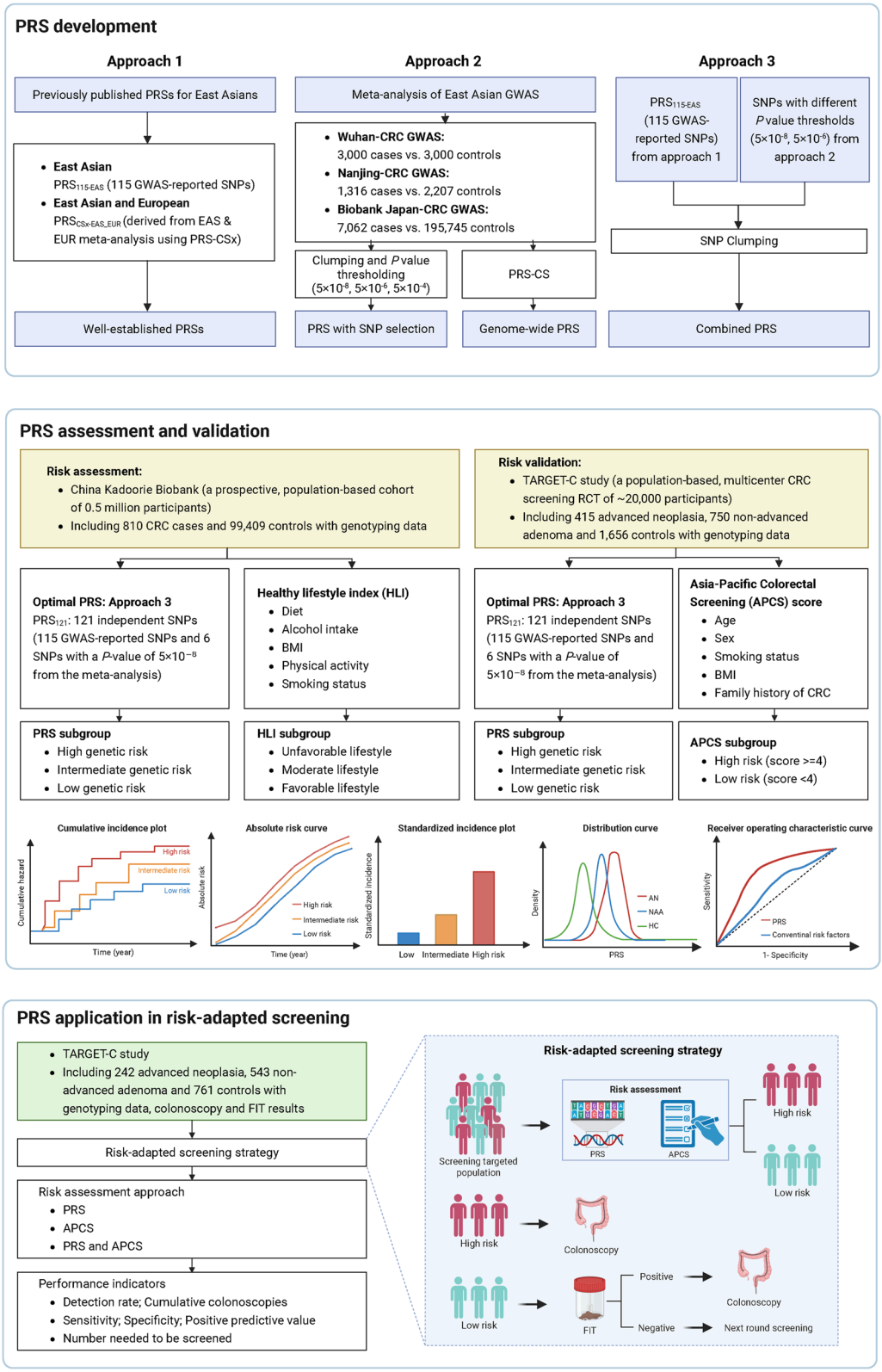



**Fig. 1** Overview of the study design. First, three PRS development approaches were used. After assessment in the CKB cohort, the optimal PRS model was selected and validated in the TARGET-C screening trial. Finally, the clinical utility of PRS-based risk-adapted CRC screening strategy was assessed. (Created in https://BioRender.com)

the participants provided written informed consent. The studies were conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Boards of Nanjing Medical University and Wuhan University. Summary statistics from the Wuhan and Nanjing CRC GWAS are publicly available [14, 15].

To increase statistical power and enhance the representation of East Asian populations, we also incorporated GWAS summary statistics for CRC from the BioBank Japan Project, which included 7,062 cases and 195,745 controls. The details have been published previously [16].

#### *Prospective cohort for PRS assessment*

The predictive performance of multiple candidate PRSs constructed using different approaches was prospectively evaluated in the China Kadoorie Biobank (CKB), a large population-based cohort study in China. The details of the study design have been reported previously [17]. In brief, a total of 512,714 adults aged 30–79 years were recruited from 10 geographically diverse areas in China between 2004 and 2008. Trained staff used electronic questionnaires to collect information and performed physical examinations and blood collection. This study was approved by the Ethical Review Committee of the Chinese Center for Disease Control and Prevention (Beijing, China) and the Oxford Tropical Research Ethics Committee of the University of Oxford (Oxford, United Kingdom). All the participants provided written informed consent. In this study, a total of 100,639 CKB participants were selected for genotyping and passed the quality control. Detailed procedures have been described elsewhere [18]. We excluded participants diagnosed with cancer at baseline; thus, 100,219 participants were included in the analysis.

#### *Population-based screening trial for PRS validation and application*

The TARGET-C study (Comparative Evaluation of Novel Screening Strategies for Colorectal Cancer Screening) is a multicenter randomized controlled trial conducted in six study centers across five provinces in China (Jiangsu, Zhejiang, Anhui, Hunan, and Yunnan). The trial aimed to compare the effectiveness and resource utilization of three CRC screening strategies: colonoscopy, fecal immunochemical testing (FIT), and risk-adapted screening. Eligible participants were adults aged 50–74 years recruited from community-based screening programs. Individuals with a prior diagnosis of CRC, severe comorbidities preventing colonoscopy, or other contraindications to screening were excluded. After providing written informed consent, participants were randomized in a 1:2:2 ratio to one-time colonoscopy ($n = 3,937$), annual FIT screening ($n = 7,858$), or a risk-adapted screening strategy ($n = 7,787$).

Participant recruitment began in May 2018, and four screening rounds (baseline and three annual follow-up rounds) were completed by October 2023. The primary outcome of the trial was the detection rate of advanced colorectal neoplasms, defined as CRC or advanced precancerous lesions. Secondary outcomes included detection of any colorectal neoplasm, screening participation, colonoscopy demand, and screening-related costs. The TARGET-C trial was prospectively registered in the Chinese Clinical Trial Registry (ChiCTR1800015506) on April 4, 2018. A detailed description of the TARGET-C study design and protocol has been published previously [19] and is provided in the Additional file 1. This study was approved by the Ethics Committees of the National Cancer Center, Chinese Academy of Medical Sciences, and Peking Union Medical College. All the participants provided written informed consent. We included participants who underwent colonoscopy during the baseline screening phase and who provided qualified plasma samples ($N = 2,821$). Detailed quality control procedures are described in the Additional file 2. The genotyping data from the TARGET-C study have been deposited in a publicly accessible repository [20].

### Outcomes

In the CKB cohort, all newly diagnosed CRC cases were identified using the 10th revision of the International Classification of Diseases (codes C18–C20). In the TARGET-C trial, the outcomes were classified as (i) CRC, (ii) advanced adenoma (AA), (iii) non-advanced adenoma (NAA), and (iv) healthy control (HC). Both CRC and AA are considered advanced neoplasias (AN). The specific definitions for each outcome are provided in the Additional file 2.

### Approaches to construct the PRS

#### *Approach 1: PRS based on the GWAS-identified CRC risk single nucleotide polymorphisms (SNPs)*

We used the $PRS_{115\text{-}EAS}$ and $PRS_{CSx\text{-}EAS\_EUR}$ as well-established EAS PRSs for subsequent assessment [11, 21]. The selection basis is provided in the Additional file 2. We calculated the PRS based on the following formula: $PRS = \sum_{i=1}^{n} \beta_i SNP_i$, where n indicates the number of SNPs, $SNP_i$ indicates the number of risk alleles for the i-th SNP (0, 1, or 2), and βi indicates the effect size carried by the risk allele.

#### *Approach 2: PRS with SNP selection based on the GWAS meta-analysis*

We performed a meta-analysis based on the summary statistics of the Nanjing, Wuhan, and Japanese GWAS using the METAL software [22]. SNPs with substantial heterogeneity ($P_{heterogeneity} < 0.001$) or those that were not present in any of the GWAS were excluded. A total of 2.2

million SNPs remained for meta-analysis. FUMA GWAS was used to perform functional annotation of the GWAS results [23]. We developed several candidate PRSs using clumping and *P*-value thresholding, and PRS-CS [24]. We set the linkage disequilibrium (LD) $r^2$ to 0.01, with different *P* value thresholds ($5\times10^{-8}$, $5\times10^{-6}$, $5\times10^{-4}$). For the PRS-CS, we restricted our study to HapMap3 SNPs and used a precalculated LD reference panel of East Asians.

#### *Approach 3: combined PRS using SNPs identified by the GWAS meta-analysis and well-established PRS*

We constructed a combined PRS by integrating variants from the $PRS_{115\text{-}EAS}$ with additional SNPs identified through a GWAS meta-analysis. To ensure that the newly added variants provided independent information beyond the established PRS, a two-step SNP selection procedure was applied. First, within each prespecified *P*-value threshold ($5\times10^{-8}$ and $5\times10^{-6}$), lead SNPs were identified using LD clumping ($r^2<0.01$ within a 1 Mb window). Second, to avoid overlap with the SNPs included in PRS115-EAS, SNPs showing LD with any $PRS_{115\text{-}EAS}$ SNP ($r^2\geq0.1$) were excluded. For the 115 SNPs from $PRS_{115\text{-}EAS}$, we used the effect sizes reported in the published model from approach 1. This strategy allowed us to enrich the PRS with additional meta-analysis-derived SNPs while maintaining a cost-effective design by avoiding a genome-wide SNP selection strategy, such as PRS-CS.

### Construction of a healthy lifestyle index using the CKB

A modified healthy lifestyle index (HLI) based on the World Cancer Research Fund recommendations was applied, which includes diet, alcohol intake, smoking status, physical activity, and body weight measures [25]. The detailed definitions of these components and scoring systems are provided in Additional file 3: Table S1.

### Calculation of a modified Asia–Pacific Colorectal Screening (APCS) score in the TARGET-C

We used a modified APCS score, which enables risk stratification according to age, sex, FH of CRC, smoking status, and body mass index (BMI) [26]. The scoring details are provided in the Additional file 2. Participants with a score ≥ 4 were defined as high risk.

### Risk-adapted screening strategy in the TARGET-C trial

Risk-adapted screening was implemented using four strategies: the PRS-based strategy, the APCS-based strategy, the dual high-risk strategy, and the any high-risk strategy. For the PRS-based strategy, participants with a PRS in the top 20% of the distribution were classified as high risk, while for the APCS-based strategy, high risk was defined as an APCS score ≥ 4. For the dual high-risk strategy, participants were classified as high risk only if they had both an APCS score ≥ 4 and a PRS in the top 20%. For the any high-risk strategy, high risk was defined as having either an APCS score ≥ 4 or a PRS in the top 20%.

In all risk-adapted strategies, individuals at high risk were referred directly for colonoscopy. Low-risk participants underwent the FIT, and those with positive FIT results were assigned to undergo a colonoscopy (Fig. 1). The FIT product used in our study was the FIT OC-SENSOR (Eiken Chemical Co, Tokyo, Japan). Following the manufacturer's recommendation, a cut-off value of ≥ 100 ng Hb/mL was used to define a positive result. Details of the colonoscopies and FIT tests are provided in the Additional file 2.

### Statistical analysis

In the CKB, we applied the Fine–Gray competing risk model to estimate the impact of the PRS on CRC [27]. The results are reported as the hazard ratio (HR) and 95% confidence interval (CI). Participants were followed-up from the date of recruitment until the date of CRC diagnosis, death, loss to follow-up, or December 31, 2018, whichever came first. The models were adjusted for sex, the top ten genetic principal components, and the HLI in the analysis of genetic risk. The discriminatory ability of the PRSs was evaluated using the C-index and HRs.

We further compared the PRS distribution between the HC and CRC groups. Adjusted HRs across the PRS quintiles were calculated. We assessed the combined effects of genetic and lifestyle factors on CRC risk. The PRS was categorized into low (bottom 20%), intermediate (20–80%), and high genetic risk (top 20%). Lifestyle was classified as unfavorable (bottom tertile), moderate (middle tertile), or favorable (top tertile) according to the HLI tertiles. The log-rank test was used to evaluate the difference in cumulative incidence (one minus the Kaplan–Meier estimate) stratified by different levels of PRS or HLI. The absolute risk was calculated using the Aalen–Johansen estimator [28]. Age-standardized incidence rates were calculated using data from the 2020 Chinese Population Census [29]. We also analyzed the interaction between genetic risk and lifestyle.

In TARGET-C, we used multivariable logistic regression to assess the impact of the PRS on different neoplasms, adjusting for age, sex, the top ten genetic principal components, and the FH of CRC. The results are reported as odds ratios (ORs) and 95% CIs. We also calculated the adjusted ORs across the PRS quintiles. Further, we assessed the combined effect of genetic risk and APCS on CRC risk using multivariate logistic regression models. In addition, we investigated the combined effects and interactions between genetic risk and FH.

The predictive performance of the PRS was compared with conventional risk factors, including age, sex, FH,

BMI, and the APCS score. We further evaluated the predictive performances of different combinations of these factors. Indicators included the area under the curve (AUC), integrated discrimination improvement (IDI), and net reclassification index. To account for potential confounding by established risk factors, we also estimated the partial AUCs. The risk-adapted screening strategy was compared to the conventional FIT-only screening strategy (. Indicators included the detection rate, sensitivity, specificity, positive predictive value (PPV), cumulative required colonoscopies and number needed to be screened (NNS).

To assess the generalizability of the findings, we compared the baseline characteristics between the CKB cohort and TARGET-C trial, and between TARGET-C participants with and without available FIT results. Standardized mean differences (SMDs) were calculated for the demographic, lifestyle, and genetic variables.

PLINK (version 1.9) [30] and the PRS-CS (version 1.0.0) [24] were used to construct the PRS. Statistical analyses were performed using R software (version 4.3) [31], and a two-sided *P*-value less than 0.05 was considered to indicate statistical significance.

## Results

### Three approaches to develop the PRS and the assessment results in the CKB cohort

To construct an optimal PRS for CRC risk stratification in the EAS population, we developed PRS models using three approaches (Fig. 1). In approach 2, the combined EAS GWAS datasets of CRC comprised 12,621 cases and 164,593 controls. We identified 15 independent SNPs that were significantly associated with CRC risk beyond genome-wide significance ($P < 5 \times 10^{-8}$; Fig. 2a). A QQ plot of the GWAS meta-analysis is shown in Additional file 4: Fig S1. All these SNPs were located within 1 Mb of well-identified regions reported by previous GWAS. The annotation of the significant SNPs is shown in Fig. 2b. A total of 100,219 participants were enrolled in the CKB, including 810 incident CRC cases, during a median follow-up of 14.0 years (Fig. 1). The baseline characteristics of the participants are summarized in Additional file 3: Table S2. The performance metrics of the different PRSs models in the CKB cohort are listed in Table 1. $PRS_{115\text{-}EAS}$ demonstrated a moderate discriminatory ability for CRC (C-index = 0.596; HR = 1.41), comparable to that of $PRS_{CSx\text{-}EAS\_EUR}$ (C-index = 0.598; HR = 1.41). For the C + T method, three different parameter settings were evaluated (LD $r^2 = 0.01$, *P* values = $5 \times 10^{-8}$, $5 \times 10^{-6}$, $5 \times 10^{-4}$), and the 5e-6 (0.01) parameter yielded the best results (C-index = 0.562; HR = 1.25). PRS-CS (auto) showed improved prediction (C-index = 0.572; HR = 1.45). Notably, the combination of $PRS_{115\text{-}EAS}$ with six SNPs from the GWAS meta-analysis that met the *P*-value threshold of $5 \times 10^{-8}$, which consisted of 121 SNPs ($PRS_{121}$), outperformed the other PRS models. $PRS_{121}$ achieved the highest discriminatory ability (C-index = 0.602; HR = 1.42; Table 1). Therefore, it was selected as the optimal PRS for further analysis. A summary of the SNPs used to construct $PRS_{121}$ is shown in Additional file 3: Table S3.

### Joint impact of PRS121 and lifestyle on CRC in the CKB cohort

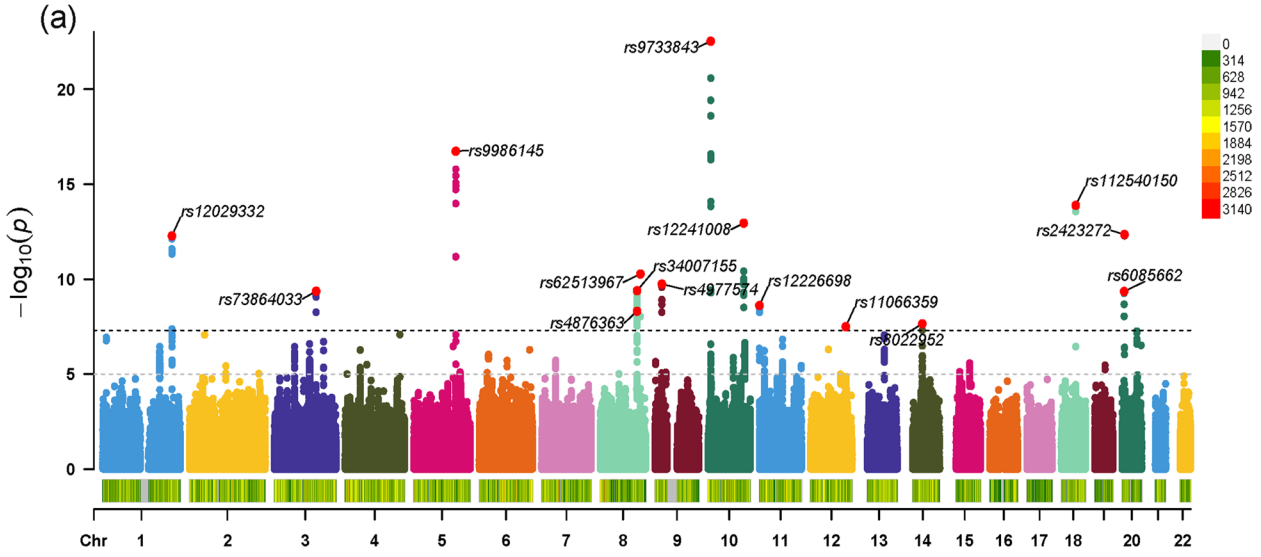
Patients with CRC tended to have higher PRS values than controls (*P* = 1.45E-22; Fig. 3a). We observed a dose–response effect across the PRS quintiles, with the risk increasing progressively from the lowest to the highest quintile. The highest quintile (top 20%) showed a 2.69-fold increase in CRC risk compared with the lowest quintile (Fig. 3b).

The cumulative incidence curves showed significant differences in CRC risk across the distinct groups ($P < 0.0001$; Fig. 3c, Additional file 4: Fig S2). Using individuals with low genetic risk and favorable lifestyle as the reference group, the highest risk was observed in participants with high genetic risk and an unfavorable lifestyle (HR = 3.32, 95% CI: 1.80–6.11), followed by those with a high genetic risk and moderate lifestyle (HR = 2.24, 95% CI: 1.23–4.05) (Fig. 3d). The cumulative risk curves for developing CRC stratified by different PRS and HLI categories are presented in Additional file 4: Fig S3 and S4.

Age-specific analyses showed clear separation of the absolute risk trajectories across different PRS and lifestyle groups. By the age of 75 years, individuals with a high PRS and an unfavorable lifestyle reached an absolute risk of approximately 3% (Fig. 3e). Although no statistically significant multiplicative or additive interaction was detected between genetic risk and lifestyle (Additional file 3: Table S4), the clinical benefit of a favorable lifestyle was more pronounced among individuals with high genetic risk. Specifically, the difference in standardized incidence rates between unfavorable and favorable lifestyles was 135.8 events per 100,000 person-years in the high genetic risk group, compared to only 20.9 events per 100,000 person-years in the low genetic risk group (Fig. 3f).

### Joint impact of PRS121 and APCS in the TARGET-C screening cohort

A total of 2,821 participants were included from a population-based CRC screening trial (TARGET-C), including 415 patients with advanced neoplasia (AN; 31 CRC and 384 AA), 750 patients with NAA, and 1,656 HCs (Fig. 1). The baseline characteristics of the participants are summarized in Additional file 3: Table S5. The CKB and TARGET-C cohorts showed small-to-moderate differences in most demographic and lifestyle variables at baseline. The SMD for age was larger




(b)

| RsID | Chr | Position | Effect allele | Other allele | MAF | Nearest gene | Function |
|---|---|---|---|---|---|---|---|
| rs12029332 | 1 | 222148822 | T | C | 0.2054 | RP11-815M8.1:RP11-400N13.2 | ncRNA_intronic |
| rs73864033 | 3 | 133736982 | T | C | 0.3641 | SLCO2A1 | intronic |
| rs9986145 | 5 | 134453390 | A | G | 0.2371 | C5orf66 | intronic |
| rs4876363 | 8 | 117636329 | A | G | 0.4087 | EIF3H | intergenic |
| rs34007155 | 8 | 117671808 | A | G | 0.1151 | EIF3H | intronic |
| rs62513967 | 8 | 128395508 | T | G | 0.2490 | CASC8:RP11-382A18.2 | ncRNA_intronic |
| rs4977574 | 9 | 22098574 | A | G | 0.4692 | CDKN2B-AS1 | ncRNA_intronic |
| rs9733843 | 10 | 8718004 | A | G | 0.3581 | RNA5SP299 | intergenic |
| rs12241008 | 10 | 114280702 | T | C | 0.2996 | VTI1A | intronic |
| rs12226698 | 11 | 213740 | T | C | 0.1548 | RIC8A | intronic |
| rs11066359 | 12 | 113045654 | T | C | 0.2917 | RPH3A | intronic |
| rs8022952 | 14 | 51394480 | T | C | 0.4345 | PYGL | intronic |
| rs112540150 | 18 | 46449925 | T | C | 0.1270 | SMAD7 | intronic |
| rs6085662 | 20 | 6698372 | C | G | 0.2242 | RP5-859D4.3 | intergenic |
| rs2423272 | 20 | 7806450 | T | C | 0.3363 | RP5-836E8.1 | intergenic |

**Fig. 2** Results for identified CRC risk loci from the GWAS meta-analysis. **a** Manhattan plot of GWAS meta-analysis results, the P values (-log10) of the SNPs (y-axis) are presented according to their chromosomal positions (x-axis). Horizontal dashed line represents the genome-wide significance threshold ($P = 5 \times 10^{-8}$). Fifteen loci exceeded genome-wide significance, with the most significant SNPs labeled. **b** Summary table of the 15 genome-wide significant loci. Abbreviations: Chr, chromosome; MAF, minor allele frequency; ncRNA, non-coding RNA

(0.86), which was consistent with the age eligibility criteria of the TARGET-C screening trial (50–74 years). The $PRS_{121}$ distribution was similar between the two cohorts, with an SMD of 0.20 (Additional file 3: Table S6).

We observed that the PRS distribution followed an increasing trend across the dynamic adenoma-carcinoma sequences. Individuals with AN exhibited a significantly higher PRS than those with NAA ($P = 0.021$) and healthy controls ($P < 0.001$). The NAA group showed a higher PRS than the healthy control group ($P < 0.001$; Fig. 4a). To investigate the risk differences across various disease stages, we categorized the participants into distinct case and control groups: (i) AN and NAA versus HC, (ii) AN versus HC, (iii) AN versus NAA and HC, and (iv) AN versus NAA. The strongest association was observed between the AN and HC groups. ORs increased steadily across PRS quintiles and reached 3.66 in the top 20% of the PRS distribution. Similar trends were observed in other comparisons (Fig. 4b).

Additionally, we assessed the combined effects of different genetic risk factors and APCS levels. The highest risk was observed in the high genetic risk and high APCS group for AN vs. HC (OR = 4.12, 95% CI: 2.12–8.01; Fig. 4c). No multiplicative or additive interactions were observed between the genetic risk and APCS levels (Additional file 3: Table S7).

**Table 1** The CRC risk prediction performance metrics of PRSs derived from different approaches in the CKB cohort

| PRS method | Parameter | $N_{SNP}$ | C-index (95% CI) | HR[a] (95% CI) | *P* value | HR[b] (95%CI) | *P* value |
|---|---|---|---|---|---|---|---|
| Well-developed | $PRS_{115\text{-}EAS}$[c] | 115 | 0.596 (0.581, 0.622) | 1.40 (1.31, 1.50) | 1.80E-22 | 1.41 (1.31, 1.50) | 6.64E-23 |
| | $PRS_{CSx\text{-}EAS\_EUR}$[d] | 1,145,689 | 0.598 (0.583, 0.623) | 1.41 (1.32, 1.51) | 1.60E-22 | 1.41 (1.31, 1.51) | 4.23E-21 |
| Clumping and *P* value thresholding | *P* = 5e-4 (LD $r^2$ = 0.01) | 564 | 0.557 (0.545, 0.586) | 1.23 (1.15, 1.31) | 1.97E-09 | 1.23 (1.15, 1.32) | 5.48E-09 |
| | *P* = 5e-6 (LD $r^2$ = 0.01) | 53 | 0.562 (0.547, 0.588) | 1.25 (1.17, 1.34) | 3.12E-10 | 1.25 (1.17, 1.34) | 3.64E-10 |
| | *P* = 5e-8 (LD $r^2$ = 0.01) | 15 | 0.559 (0.542, 0.586) | 1.24 (1.16, 1.34) | 4.46E-09 | 1.24 (1.16, 1.34) | 4.95E-09 |
| PRS-CS | Auto | 417,502 | 0.572 (0.553, 0.591) | 1.30 (1.21, 1.40) | 2.59E-12 | 1.45 (1.33, 1.58) | 1.93E-17 |
| Combined | $PRS_{115}$ + 5e-6 (LD $r^2$ = 0.01) | 143 | 0.600 (0.582, 0.623) | 1.40 (1.31, 1.49) | 2.90E-23 | 1.40 (1.31, 1.50) | 2.99E-23 |
| | $PRS_{115}$ + 5e-8 (LD $r^2$ = 0.01) | 121 | 0.602 (0.582, 0.622) | 1.41 (1.32, 1.51) | 4.91E-24 | 1.42 (1.33, 1.52) | 1.30E-24 |

*Abbreviations*: *CRC* Colorectal cancer, *CKB* China Kadoorie Biobank, *PRS* Polygenic risk score, *HR* Hazard ratio, *CI* Confidence interval, *LD* Linkage disequilibrium

[a]The model includes continuous z-transformed PRS

[b]The model includes age, sex, top 10 principal components, and continuous z-transformed PRS

[c]Developing and validating polygenic risk scores for colorectal cancer risk prediction in East Asians. Int J Cancer. 2022 Nov 15;151(10):1726–1736

[d]Risk assessment for colorectal cancer via polygenic risk score and lifestyle exposure: a large-scale association study of East Asian and European populations. Genome Med. 2023 Jan 24;15(1):4

### Discriminative performance of PRS121 and conventional risk factors in the TARGET-C screening cohort

The PRS demonstrated the highest discriminative ability among the conventional risk factors, including age, sex, BMI, FH, and APCS. The PRS consistently achieved the highest AUC and partial AUC among all comparisons (AUC: 0.629 for AN vs. HC, 0.591 for AN + NAA vs. HC, 0.598 for AN vs. NAA + HC, and 0.554 for AN vs. NAA; partial AUC: 0.624 for AN vs. HC, 0.590 for AN + NAA vs. HC, 0.593 for AN vs. NAA + HC, and 0.550 for AN vs. NAA; Fig. 5a, Additional file 3: Table S8).

For different risk factor combinations, incorporating PRS consistently improved the discriminative ability across all comparisons. For AN vs. HC, the AUC increased from 0.558 (age + sex) to 0.599 (adding BMI), 0.612 (adding FH of CRC), 0.626 (adding smoking), and 0.679 (adding the PRS). Similar results were observed in other comparisons (Fig. 5b). The full model (age + sex + BMI + FH + smoking + PRS) showed the best discriminative ability for all the outcomes. For AN vs. HC, simpler models had negative IDI values (e.g., age + sex: IDI = −0.059, 95% CI: −0.074 to −0.045; $P < 0.001$) and NRI values (continuous NRI = −0.447, 95% CI: −0.582 to −0.312; categorical NRI = −0.172, 95% CI: −0.246 to −0.098) compared with the full model. Similar trends were observed in other groups (Additional file 3: Table S9–S12).

### Evaluation of the screening yield of traditional and novel risk-adapted screening strategies

Among the 2,821 participants, 1,275 had no FIT results. Therefore, the performance evaluation study of the risk-adapted screening strategy ultimately included 1,546 participants (Fig. 1). Comparisons between individuals with and without available FIT results showed modest differences in demographic and lifestyle characteristics (SMDs 0.12–0.27). $PRS_{121}$ distributions were nearly identical between the two groups (SMD = 0.00, Additional file 3: Table S13).

Figure 6 compares the screening performance across different strategies. In terms of AN detection, all risk-adapted strategies demonstrated higher detection rates than the FIT-only strategy. The detection rate increased from 13.1% with FIT-only screening to 16.7% with the PRS-based strategy ($P = 0.024$), and further to 21.0% with the any high-risk strategy integrating both PRS and APCS ($P = 2.72E\text{-}06$).

Regarding diagnostic accuracy, the any high-risk strategy achieved the highest sensitivity (87.2%), substantially exceeding that of the FIT-only strategy (54.1%, $P = 3.10E\text{-}15$). In contrast, specificity and positive predictive value (PPV) were highest in the FIT-only strategy (73.9% and 39.7%, respectively) and decreased in risk-adapted strategies, reflecting the expected trade-off between sensitivity and specificity.

In terms of resource utilization, the PRS-based strategy required more colonoscopies than FIT-only screening (462 *vs.* 330), while the any high-risk strategy required the greatest number (687). However, colonoscopy efficiency remained comparable across strategies, with the number of colonoscopies required to detect one AN being 2.52 for FIT-only screening, 2.75 for the PRS-based strategy, and 3.26 for the any high-risk strategy.

Importantly, the PRS-based risk-adapted strategy identified a substantial proportion of AN cases that would have been missed by FIT-only screening. Specifically, 35 of 166 detected AN cases (21.1%) were FIT-negative but classified as high risk by PRS, indicating that over one-fifth of AN cases were uniquely identified through genetic risk stratification rather than FIT positivity. When PRS and APCS were combined, this proportion increased to 80 of 211 cases (37.9%), demonstrating that integrated risk stratification substantially improves detection of FIT-negative lesions. These findings are further illustrated in Fig. 7.

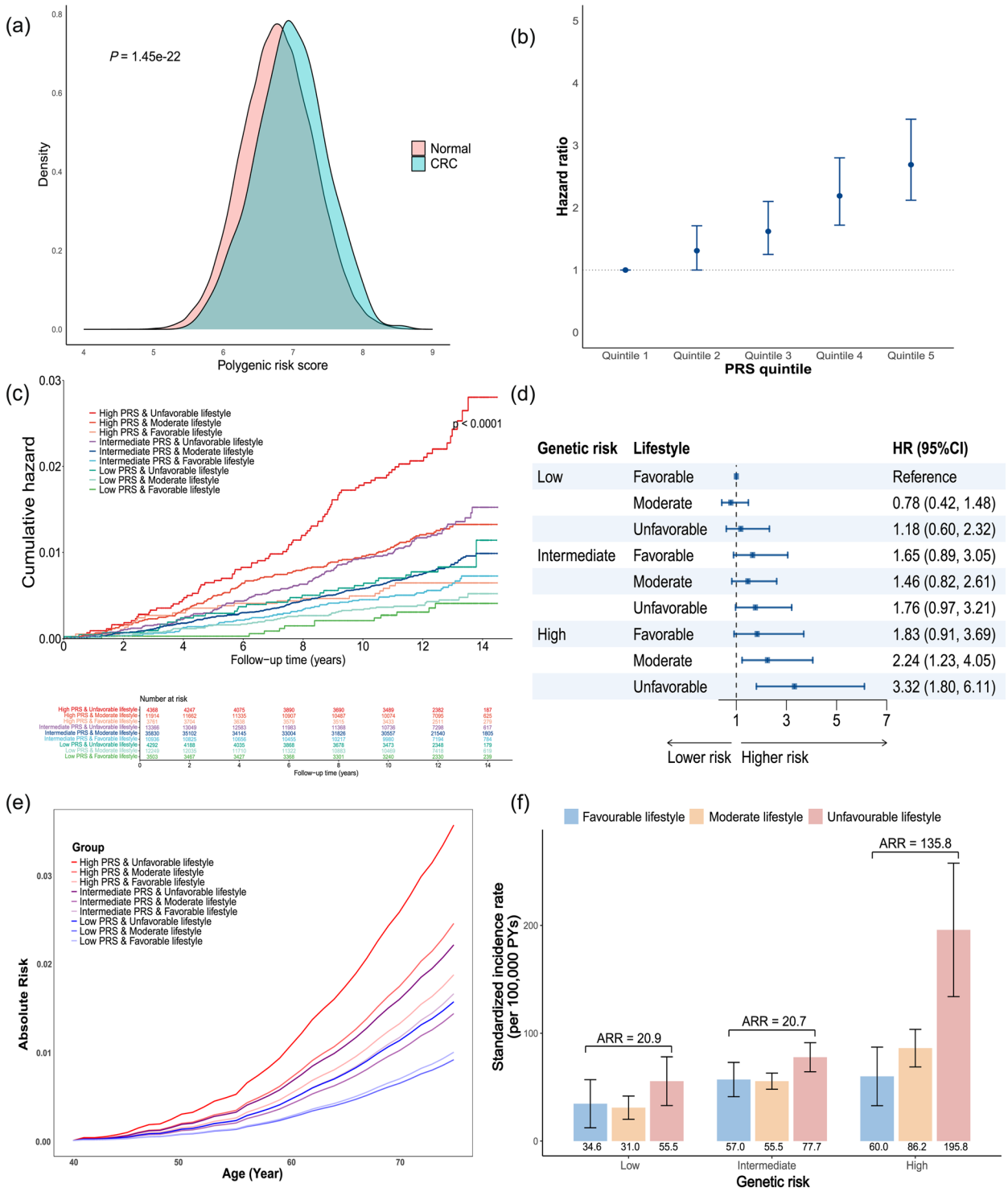
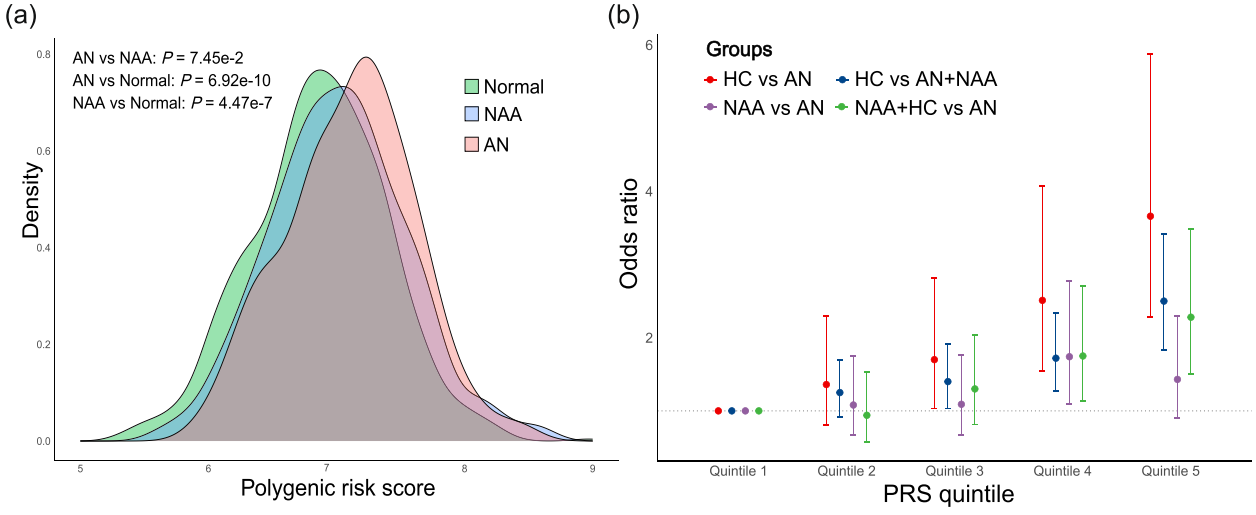



**Fig. 3** Joint associations of PRS and lifestyles with CRC risk in the CKB cohort. **a** Distribution of $PRS_{121}$ values between CRC cases and controls. **b** Dose–response relationship between PRS quintiles and CRC HRs. **c** Cumulative incidence of CRC defined by different levels of PRS and lifestyle score. **d** Forest plot of CRC risk defined by different levels of PRS and lifestyle score. **e** Absolute risk defined by different levels of PRS and lifestyle score. **f** Age-standardized CRC incidence defined by different levels of PRS and lifestyle score. PRS levels were defined as low (bottom 20%), intermediate (20–80%), and high (top 20%); lifestyle categories were stratified by healthy lifestyle index tertiles. The HR and 95% CI were derived from the Fine-Gray competing risk model with the adjustment of age, sex and first 10 principal components. Abbreviations: PRS, polygenic risk score; HR, hazard ratio; 95% CI, 95% confidence interval




(c)

| Genetic risk | APCS risk | Model 1 | OR (95%CI) | Model 2 | OR2 (95%CI) |
|---|---|---|---|---|---|
| **Advanced neoplasm and NAA vs normal** | | | | | |
| Low | Low | | Reference | | Reference |
| | High | | 1.16 (0.75, 1.80) | | 1.16 (0.75, 1.80) |
| Intermediate | Low | | Reference | | 1.32 (0.94, 1.84) |
| | High | | 1.42 (1.10, 1.83) | | 1.87 (1.32, 2.67) |
| High | Low | | Reference | | 2.69 (1.78, 4.07) |
| | High | | 0.93 (0.59, 1.47) | | 2.52 (1.60, 3.96) |
| **Advanced neoplasm vs normal** | | | | | |
| Low | Low | | Reference | | Reference |
| | High | | 1.14 (0.53, 2.45) | | 1.14 (0.53, 2.45) |
| Intermediate | Low | | Reference | | 1.43 (0.81, 2.53) |
| | High | | 1.93 (1.34, 2.79) | | 2.76 (1.55, 4.90) |
| High | Low | | Reference | | 3.56 (1.91, 6.66) |
| | High | | 1.16 (0.66, 2.02) | | 4.12 (2.12, 8.01) |
| **Advanced neoplasm vs NAA and normal** | | | | | |
| Low | Low | | Reference | | Reference |
| | High | | 1.11 (0.57, 2.17) | | 1.11 (0.57, 2.17) |
| Intermediate | Low | | Reference | | 1.01 (0.60, 1.70) |
| | High | | 1.88 (1.32, 2.68) | | 1.91 (1.14, 3.20) |
| High | Low | | Reference | | 2.31 (1.32, 4.05) |
| | High | | 1.06 (0.63, 1.78) | | 2.46 (1.34, 4.49) |
| **Advanced neoplasm vs NAA** | | | | | |
| Low | Low | | Reference | | Reference |
| | High | | 0.94 (0.47, 1.87) | | 0.94 (0.47, 1.87) |
| Intermediate | Low | | Reference | | 0.99 (0.58, 1.68) |
| | High | | 1.58 (1.08, 2.31) | | 1.57 (0.92, 2.68) |
| High | Low | | Reference | | 1.24 (0.67, 2.31) |
| | High | | 1.3 (0.69, 2.48) | | 1.62 (0.83, 3.16) |



**Fig. 4** Discriminative ability of $PRS_{121}$ across the adenoma-carcinoma sequence in the TARGET-C cohort. **a** Distribution of $PRS_{121}$ among HC, NAA, and AN cases. **b** Dose–response relationship between PRS quintiles and ORs across different case–control comparisons. **c** Joint effects of genetic risk and APCS score on colorectal neoplasms risk. PRS categories were defined as low (bottom 20%), intermediate (20–80%), and high (top 20%); APCS high-risk was defined as score ≥ 4. The OR and 95% CI were derived from the logistic regression models with the adjustment of age, sex, the top ten genetic principal components and family history of CRC. Model 1 conducted subgroup analyses within different genetic risk groups based on APCS levels, using high genetic risk with low APCS level, intermediate genetic risk with low APCS level, and low genetic risk with low APCS level as the reference groups, respectively. Model 2 used only the low genetic risk with low APCS level group as the reference. Abbreviations: PRS, polygenic risk score; HC, healthy control; NAA, non-advanced adenoma; AN, advanced neoplasia; OR, odds ratio; APCS, Asia–Pacific Colorectal Screening score; CRC, colorectal cancer; 95% CI, 95% confidence interval

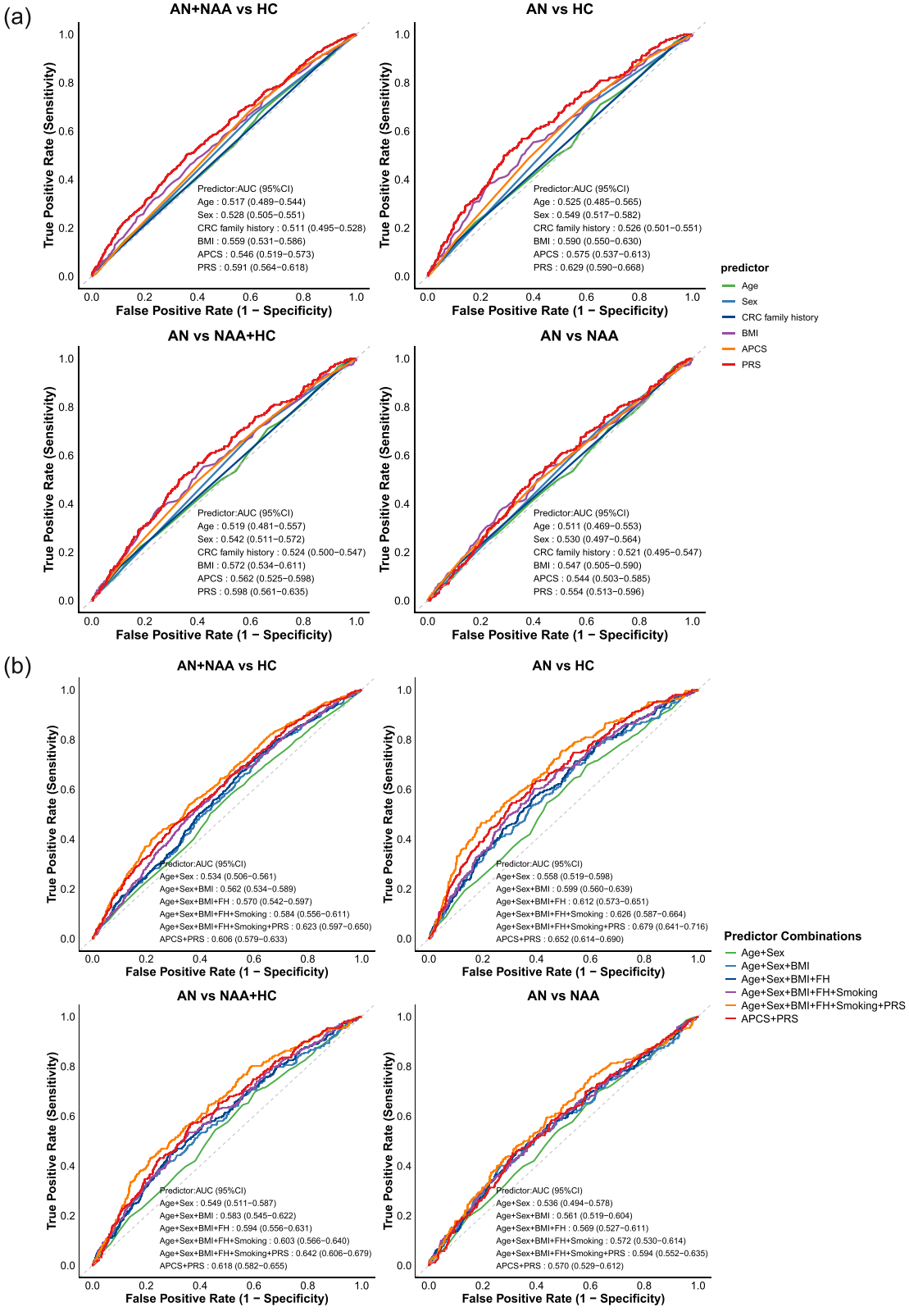



**Fig. 5** Discriminative ability of individual and combinations of $PRS_{121}$ and conventional risk factors in the TARGET-C cohort. **a** ROC curves and corresponding AUC values with 95% CIs for six individual predictors—age, sex, CRC family history, BMI, APCS, and PRS across different comparisons. **b** ROC curves for different models incorporating combinations of traditional risk factors and PRS across different comparisons. Abbreviations: HC, healthy control; NAA, non-advanced adenoma; AN, advanced neoplasia; PRS, polygenic risk score; ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence intervals; CRC, colorectal cancer; BMI, body mass index; APCS, Asia–Pacific Colorectal Screening score

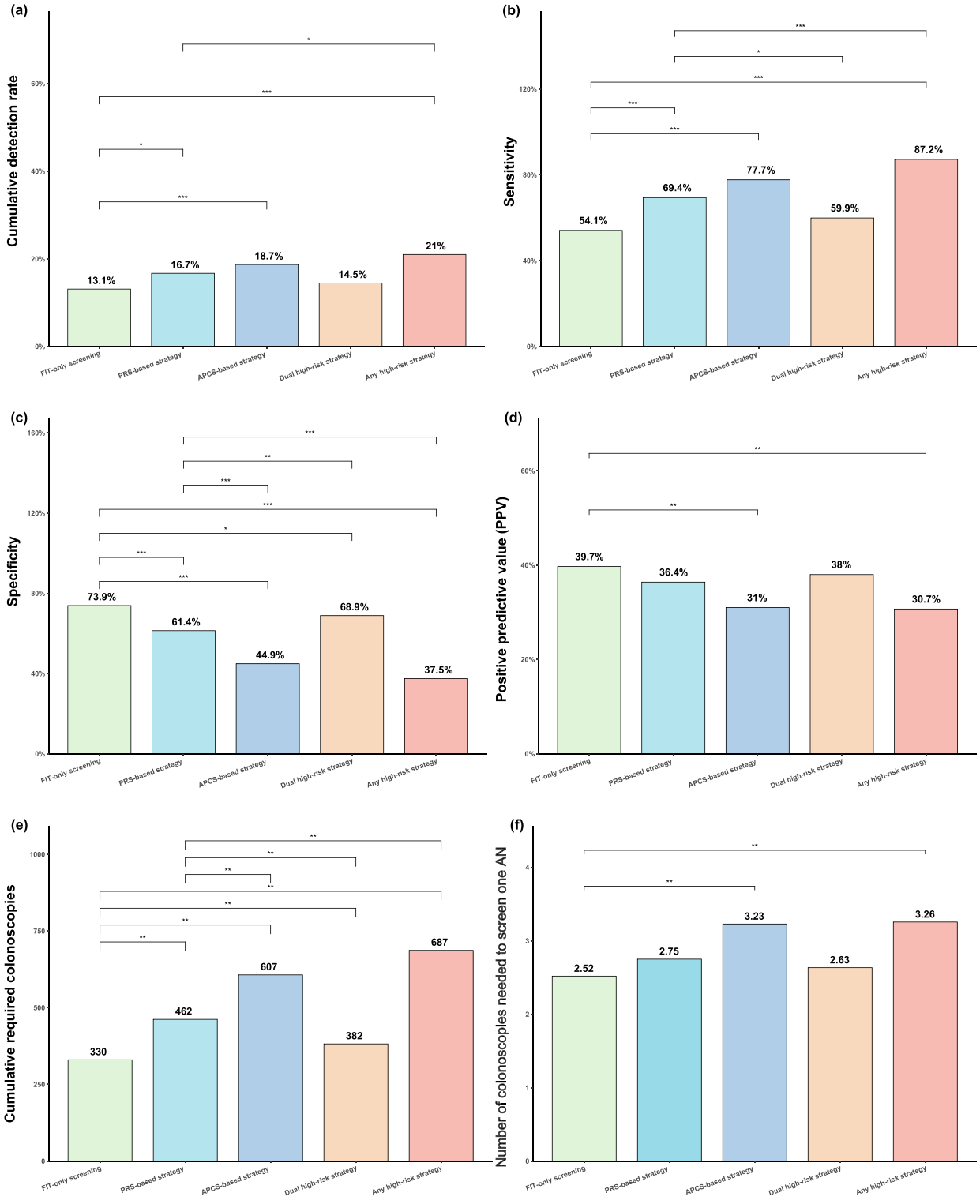



**Fig. 6** Comparison of screening performance across FIT-only and risk-adapted screening strategies. **a** Cumulative detection rate of advanced neoplasia (AN); (**b**) Sensitivity; (**c**) Specificity; (**d**) Positive predictive value (PPV); (**e**) Total number of colonoscopies required. **f** Number of colonoscopies needed to screen (NNS) one AN case. Risk-adapted strategies included PRS-based, APCS-based, dual high-risk (PRS and APCS), and any high-risk (PRS or APCS) approaches. Statistical significance was assessed using appropriate comparative tests (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). Data shown are derived from the study cohort. Abbreviations: AN, advanced neoplasia; PRS, polygenic risk score; APCS, Asia–Pacific Colorectal Screening Score

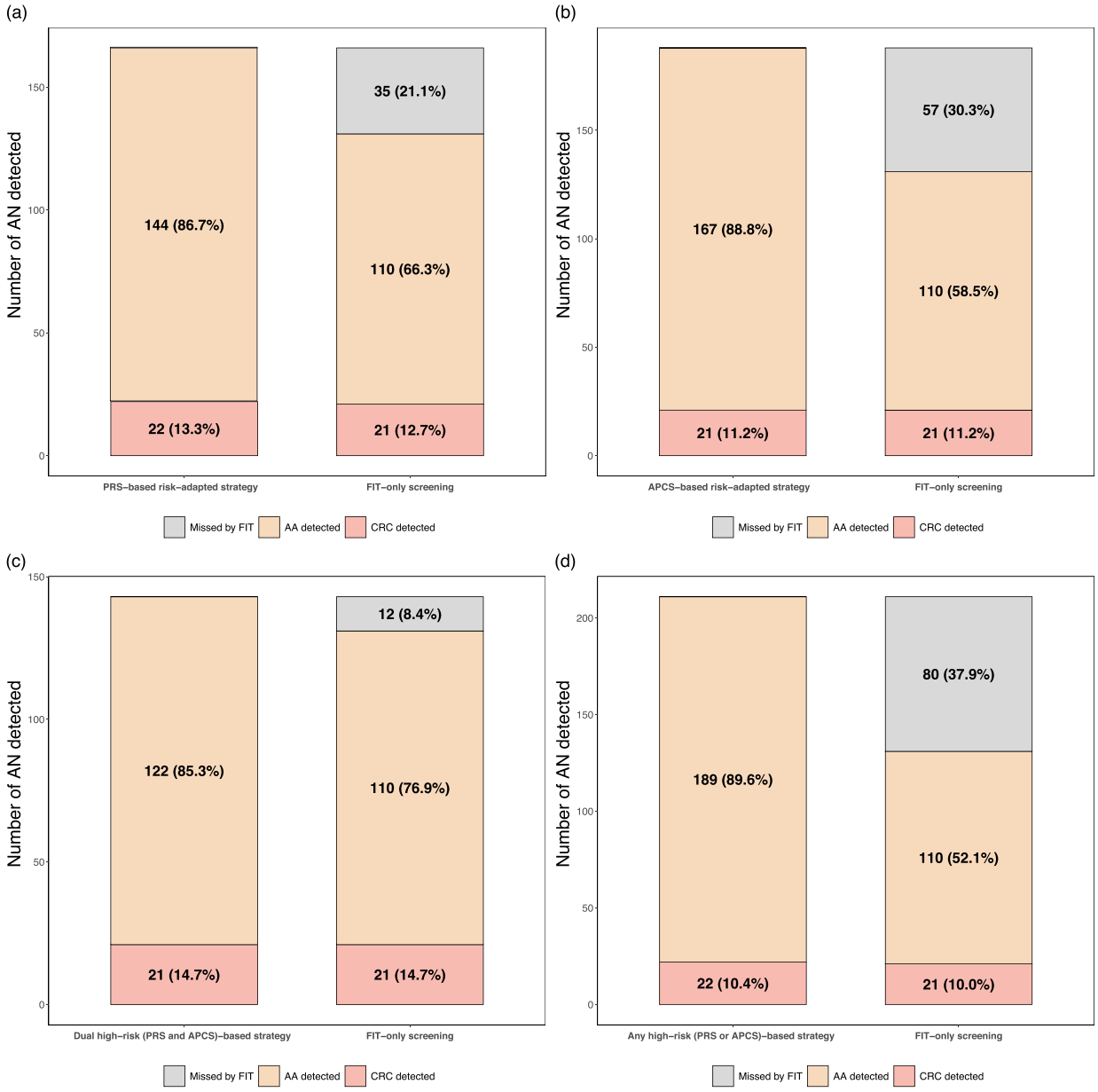



**Fig. 7** Contribution of risk-adapted screening strategies to detection of FIT-negative advanced neoplasia. Stacked bar plots showing the distribution of advanced neoplasia (AN) cases detected by FIT positivity and additional cases identified among FIT-negative individuals classified as high risk by different risk stratification strategies, including (**a**) PRS-based, (**b**) APCS-based, (**c**) dual high-risk (PRS and APCS), and (**d**) any high-risk (PRS or APCS) strategies. Abbreviations: AN, advanced neoplasia; PRS, polygenic risk score; APCS, Asia–Pacific Colorectal Screening Score

## Discussion

This study evaluated the utility of the PRS for CRC risk prediction in a large-scale prospective cohort and its utility as a risk-adapted screening strategy within a population-based screening setting. The $PRS_{121}$ demonstrated superior discriminatory performance compared with conventional risk factors and existing PRS models, highlighting the potential of integrating the PRS with established screening modalities—such as FIT and colonoscopy—within risk-adapted CRC screening strategies.

Unlike previous studies that primarily relied on cross-sectional data [10, 11], our analysis utilized the CKB, a large-scale prospective cohort, and rigorously accounted for competing risks through statistical calibration. Our results demonstrated a robust association between the PRS and CRC risk, which is consistent with previous studies [32, 33], providing compelling evidence for the utility of genetic risk stratification. The $PRS_{121}$, derived from a combination of well-established GWAS-identified SNPs and additional variants from our EAS GWAS

meta-analysis, outperformed the other models. Notably, the performance of the $PRS_{121}$ was achieved with only 121 SNPs, in contrast to genome-wide approaches such as the PRS-CS, which require substantially more variants for computation. This optimized SNP panel-based strategy offers a cost-effective alternative to genome-wide PRS methods, as it reduces the genotyping burden while maintaining robust discriminative ability.

Importantly, lifestyle factors substantially modified CRC risk across genetic risk strata. Among individuals with high genetic risk, those with favorable lifestyle had markedly lower risk compared with those with unfavorable lifestyle. These findings highlight the important role of lifestyle modification in mitigating inherited CRC risk and support its relevance for precision prevention strategies.

A key strength of our study was the assessment of the real-world clinical utility of the PRS in risk-adapted screening for CRC. To our knowledge, this is the first study to conduct such an evaluation. CRC predominantly develops via the adenoma–carcinoma sequence. Early detection and removal of precancerous lesions, particularly AA, are more effective for CRC prevention than the detection of established CRC, making it a primary objective of screening programs. In this study, we observed that PRS and FIT exhibited complementary performance in detecting advanced colorectal neoplasia: PRS demonstrated higher sensitivity, whereas FIT showed better specificity. Evidence from previous systematic reviews also indicates that the sensitivity of the FIT for detecting precancerous lesions is substantially lower than that for detecting CRC and varies depending on the positivity threshold [34]. For example, the sensitivity of FIT for detecting CRC was high (91% at a threshold of 10 μg/g), whereas its sensitivity for AA was considerably lower (40% at the same threshold) [34]. This dichotomy supports the potential utility of a risk-adapted screening approach in which the PRS could serve as an initial risk stratification tool to identify candidates for subsequent FIT testing or direct colonoscopy. Our findings demonstrate that a risk-adapted screening strategy integrating the PRS with existing screening modalities, including the FIT and colonoscopy, offers a promising approach for enhancing the CRC screening efficacy. The PRS-based risk-adapted strategy achieved a more favorable balance between sensitivity and specificity than FIT alone, with a significantly higher detection rate. This strategy may optimize the trade-offs among sensitivity, specificity, and resource utilization in population-based screening programs.

Risk-adapted screening necessitates risk assessment across the target population, which inevitably incurs additional costs. Several studies have explored the potential cost-effectiveness of PRS risk-adapted CRC screening. However, the findings are not consistent. Some studies have reported no economic health benefits from integrating the PRS into CRC population screening [35, 36]. Regarding China's organized cancer screening programs, a modeling study of the PRS-adapted screening strategy demonstrated moderate improvement in clinical benefits and cost-effectiveness [37]. Although our cross-sectional analysis suggests that PRS stratification may modestly improve screening efficiency, robust evidence from RCTs is still needed, considering ancestral inequities, societal acceptability, a potential reduction in current screening uptake, and logistical challenges. Evidence of a favorable benefit–harm balance in research settings does not necessarily guarantee similar outcomes in real-world practice.

Despite these advancements, this study has some limitations. First, the generalizability of our findings may be constrained by the predominant EAS ancestry of the study population. Validation with diverse ethnic groups is necessary to ensure broader applicability. Second, the incremental benefit of the PRS over conventional risk factors, although statistically significant, was modest for some comparisons. Further improvements in PRS models, possibly through the inclusion of rare variants or functional genomic data, may enhance predictive performance. Third, the effectiveness of PRS-based risk-adapted screening was evaluated in a cross-sectional manner. Its long-term impact on CRC incidence and mortality, compared to other strategies, should be assessed in prospective cohorts or RCTs with longer follow-up periods. Lastly, the cost-effectiveness of the PRS-based screening strategy was not evaluated in the current study, as it falls outside its scope. However, this aspect should be assessed in future studies.

## Conclusions

Our study demonstrates that PRS enables effective risk stratification for colorectal cancer and improves detection of advanced neoplasia by identifying high-risk individuals missed by FIT, supporting its utility in precision risk-adapted screening. Future research should focus on validating these findings across diverse populations and evaluating the long-term effects of PRS-based screening on CRC incidence and mortality.

### Abbreviations

| | |
|---|---|
| AA | Advanced adenoma |
| AN | Advanced neoplasias |
| APCS | Asia–Pacific Colorectal Screening |
| AUC | Area under the curve |
| BMI | Body mass index |
| CI | Confidence interval |
| CKB | China Kadoorie Biobank |
| CRC | Colorectal cancer |
| EAS | East Asian |
| FIT | Fecal immunochemical testing |
| GWAS | Genome-wide association studies |

| | |
|---|---|
| HC | Healthy control |
| HLI | Healthy lifestyle index |
| HR | Hazard ratio |
| ICD-10 | International Classification of Diseases, 10th revision |
| IDI | Integrated discrimination improvement |
| LD | Linkage disequilibrium |
| NAA | Non-advanced adenoma |
| NNS | Number of colonoscopy needed to be screened |
| PPV | Positive predictive value |
| PRS | Polygenic risk scores |
| RCT | Randomized controlled trial |
| SNP | Single nucleotide polymorphism |
| SMD | Standardized mean difference |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-026-01623-z.

Additional file 1. TARGET-C study protocol.

Additional file 2. Supplementary methods.

Additional file 3: Table S1. Construction of the healthy lifestyle index in the China Kadoorie Biobank cohort. Table S2. Baseline characteristics of the China Kadoorie Biobank cohort based on colorectal cancer status. Table S3. Summary of single nucleotide polymorphisms in the optimal polygenic risk score. Table S4. Interaction between genetic risk and healthy lifestyle index in the China Kadoorie Biobank cohort. Table S5. Baseline characteristics of the TARGET-C trial based on colorectal neoplasm status. Table S6. Baseline characteristic comparison between China Kadoorie Biobank and TARGET-C. Table S7. Interaction between genetic risk and APCS categories in the TARGET-C. Table S8. Partial AUC results with 95% CIs for six individual predictors. Table S9. Predictors and performance of different risk factor combinations (Advanced neoplasm and NAA vs normal). Table S10. Predictors and performance of different risk factor combinations (Advanced neoplasm vs normal). Table S11. Predictors and performance of different risk factor combinations (Advanced neoplasm vs NAA and normal). Table S12. Predictors and performance of different risk factor combinations (Advanced neoplasm vs NAA). Table S13. Baseline characteristic comparison between FIT available and unavailable participants in TARGET-C.

Additional file 4: Figure S1. QQ plot of GWAS meta-analysis. Figure S2. Cumulative incidence plot of colorectal cancer according to different genetic risk and lifestyle categories. Figure S3. Cumulative incidence plot of incident colorectal cancer according to healthy lifestyle index categories. Figure S4. Cumulative incidence plot of incident colorectal cancer according to genetic risk categories.

Additional file 5.

**Acknowledgements**

The authors would like to thank all the staff involved in the fieldwork of the China Kadoorie Biobank and TARGET-C study for their valuable contributions to this work. We also thank all study members who contributed to the completion of this research.

**Authors' contributions**

HC, JS, CL, JT, and JX conceptualized and designed the study. HC, JS, CL, JT, and JX participated in the statistical analysis and visualization. MW, XM, JL and MD were responsible for supervision of the project administration and quality assessment. CY, DS, PP, LY, IM, RW, YC, HD, ZC, LL, MK, YK participated in the acquisition of data and analysis and interpretation of data. HC, JS, CL, JT, and JX drafted the manuscript. All authors reviewed and edited subsequent drafts. All authors accessed the data reported in the study and approved the final version of the manuscript.

**Funding**

This work was supported by the Beijing Research Ward Excellence Program (BRWEP2024W034010101), the National High Level Hospital Clinical Research Funding (2025-LYZX-C-B03), the Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education (2024101), the Chinese Academy of Medical Science Innovation Fund for Medical Science (2022-I2M-1–003), the National Natural Science Foundation of China (82173606; 82273726; 82473705; 82388102), the Science and Technology Planning Project of Tibet Autonomous Region (XZ202501JD0021), the PUMCH Talent Development and Support Program, Category B (UGG06641). The CKB baseline survey and the first re-survey were supported by the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up has been supported by Wellcome grants to Oxford University (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z) and grants (2016YFC0900500) from the National Key R&D Program of China, the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2023ZD0510100), and the National Natural Science Foundation of China (82192900, 81390540, 91846303, 81941018), and Chinese Ministry of Science and Technology (2011BAI09B01). The funders had no role in the study design and conduct; data collection, management, analysis, and interpretation; manuscript preparation, review, or approval; and the decision to submit the manuscript for publication.

**Data availability**

Details of how to access and details of the CKB data release are available from www. ckbiobank.org/site/Data + Access. As stated in the access policy, the CKB study group must maintain the integrity of the database for future use and regulate data access to comply with prior conditions agreed with the Chinese government. Data security is an integral part of CKB study protocols. Data can be released outside the CKB research group only with appropriate security safeguards. GWAS summary statistics of BBJ is available from the website (http://jenger.riken.jp/result). The TARGET-C genotyping dataset generated in this study has been deposited in the Science Data Bank repository (https://www.scidb.cn/) and is publicly available at: https://cstr.cn/31253.11.sciencedb.28745. The GWAS summary statistics from the Wuhan CRC GWAS and Nanjing CRC GWAS used for PRS construction have also been deposited in the Science Data Bank and are available at: https://cstr.cn/31253.11.sciencedb.28736 and https://cstr.cn/31253.11.sciencedb.28742. The PRS developeld in this study (PRS 121) has been submitted to the Polygenic Score (PGS) Catalog (https://www.pgscatalog.org/) and the accession number is PGS005402. Meanwhile, detailed parameters of PRS 121 is also provided in Additional file 3: Table S3.

## Declarations

**Ethics approval and consent to participate**

All participants in the included studies provided written informed consent. The Nanjing GWAS and Wuhan GWAS were approved by the Institutional Review Board of Nanjing Medical University and Wuhan University, respectively. The CKB cohort study was approved by the Ethical Review Committee of the Chinese Center for Disease Control and Prevention (Beijing, China) and the Oxford Tropical Research Ethics Committee, University of Oxford (Oxford, United Kingdom). The TARGET-C study was approved by the Ethics Committee of the National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College, and prospectively registered in the Chinese Clinical Trial Registry (ChiCTR1800015506). The research was conducted in accordance with the principles of the Declaration of Helsinki.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no competing interests.

**Author details**

[1]Center for Prevention and Early Intervention, National Infrastructures for Translational Medicine, Institute of Clinical Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China
[2]CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China
[3]Department of Cancer Epidemiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[4]Department of Epidemiology and Biostatistics, School of Public Health; State Key Laboratory of Metabolism and Regulation in Complex Organisms, TaiKang Center for Life and Medical Sciences, Wuhan University, Wuhan, China
[5]Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China
[6]Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing, China
[7]Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing, China
[8]Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China
[9]Medical Research Council Population Health Research Unit at the University of Oxford, Oxford, UK
[10]Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Oxford, UK
[11]Laboratory of Complex Trait Genomics, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan
[12]Department of Environmental Genomics, Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing, China
[13]State Key Laboratory of Vascular Homeostasis and Remodeling, Peking University, Beijing, China
[14]Department of Epidemiology and Biostatistics, School of Public Health, Research Center of Public Health, Renmin Hospital of Wuhan University, Wuhan, China
[15]Department of Gastrointestinal Oncology, Zhongnan Hospital of Wuhan University, TaiKang Center for Life and Medical Sciences, Wuhan University, Wuhan, China

Received: 27 June 2025 / Accepted: 6 March 2026
Published online: 19 March 2026

## 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. https://doi.org/10.3322/caac.21834.
2. Ding H, Lin J, Xu Z, Wang HHX, Huang L, Huang J, et al. The association between organised colorectal cancer screening strategies and reduction of its related mortality: a systematic review and meta-analysis. BMC Cancer. 2024;24:365. https://doi.org/10.1186/s12885-024-12054-7.
3. Cardoso R, Guo F, Heisser T, Hackl M, Ihle P, De Schutter H, et al. Colorectal cancer incidence, mortality, and stage distribution in European countries in the colorectal cancer screening era: an international population-based study. Lancet Oncol. 2021;22:1002–13. https://doi.org/10.1016/S1470-2045(21)00199-6.
4. Shaukat A, Levin TR. Current and future colorectal cancer screening strategies. Nat Rev Gastroenterol Hepatol. 2022;19:521–31. https://doi.org/10.1038/s41575-022-00612-y.
5. Kastrinos F, Kupfer SS, Gupta S. Colorectal cancer risk assessment and precision approaches to screening: brave new world or worlds apart? Gastroenterology. 2023;164:812–27. https://doi.org/10.1053/j.gastro.2023.02.021.
6. Chen H, Shi J, Lu M, Li Y, Du L, Liao X, et al. Comparison of colonoscopy, fecal immunochemical test, and risk-adapted approach in a colorectal cancer screening trial (TARGET-C). Clin Gastroenterol Hepatol. 2023;21:808–18. https://doi.org/10.1016/j.cgh.2022.08.003.
7. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, et al. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet. 2019;51:76–87. https://doi.org/10.1038/s41588-018-0286-6.
8. Fernandez-Rozadilla C, Timofeeva M, Chen Z, Law P, Thomas M, Schmit S, et al. Deciphering colorectal cancer genetics through multi-omic analysis of 100,204 cases and 154,587 controls of European and east Asian ancestries. Nat Genet. 2023;55:89–99. https://doi.org/10.1038/s41588-022-01222-9.
9. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15:2759–72. https://doi.org/10.1038/s41596-020-0353-1.
10. Tian J, Zhang M, Zhang F, Gao K, Lu Z, Cai Y, et al. Developing an optimal stratification model for colorectal cancer screening and reducing racial disparities in multi-center population-based studies. Genome Med. 2024;16:81. https://doi.org/10.1186/s13073-024-01355-y.
11. Xin J, Du M, Gu D, Jiang K, Wang M, Jin M, et al. Risk assessment for colorectal cancer via polygenic risk score and lifestyle exposure: a large-scale association study of East Asian and European populations. Genome Med. 2023;15:4. https://doi.org/10.1186/s13073-023-01156-9.
12. Xin J, Du M, Gu D, Ge Y, Li S, Chu H, et al. Combinations of single nucleotide polymorphisms identified in genome-wide association studies determine risk for colorectal cancer. Int J Cancer. 2019;145:2661–9. https://doi.org/10.1002/ijc.32267.
13. Lu Z, Chen C, Zhang H, Li B, Liu Y, Guo J, et al. Characterization of *cis*-regulatory elements and functional variants in colorectal cancer using epigenomics and CRISPRi screenings. Nat Cancer. 2025;6:1777–99. https://doi.org/10.1038/s43018-025-01031-z.
14. Tian J, Miao X. Wuhan GWAS summary statistics. https://www.scidb.cn/. Science Data Bank; 2026. https://doi.org/10.57760/sciencedb.28736. Cited 2026 Mar 4.
15. Xin J, Wang M. Nanjing GWAS summary statistics. https://www.scidb.cn/. Science Data Bank; 2026. https://doi.org/10.57760/sciencedb.28742. Cited 2026 Mar 4.
16. Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, et al. Large scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. Nat Genet. 2020;52:669–79. https://doi.org/10.1038/s41588-020-0640-3.
17. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol. 2011;40:1652–66. https://doi.org/10.1093/ije/dyr120.
18. Walters RG, Millwood IY, Lin K, Schmidt Valle D, McDonnell P, Hacker A, et al. Genotyping and population characteristics of the China Kadoorie Biobank. Cell Genom. 2023;3:100361. https://doi.org/10.1016/j.xgen.2023.100361.
19. Chen H, Li N, Shi J, Ren J, Liu C, Zhang Y, et al. Comparative evaluation of novel screening strategies for colorectal cancer screening in China (TARGET-C): a study protocol for a multicentre randomised controlled trial. BMJ Open. 2019;9:e025935. https://doi.org/10.1136/bmjopen-2018-025935.
20. Chen H, Luo C, Dai M. TARGET-C genotyping data. https://www.scidb.cn/. Science Data Bank; 2026. https://doi.org/10.57760/sciencedb.28745. Cited 2026 Mar 4.
21. Ping J, Yang Y, Wen W, Kweon S, Matsuda K, Jia W, et al. Developing and validating polygenic risk scores for colorectal cancer risk prediction in East Asians. Int J Cancer. 2022;151:1726–36. https://doi.org/10.1002/ijc.34194.
22. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26:2190–1. https://doi.org/10.1093/bioinformatics/btq340.
23. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826. https://doi.org/10.1038/s41467-017-01261-5.
24. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019;10:1776. https://doi.org/10.1038/s41467-019-09718-5.
25. Shams-White MM, Brockton NT, Mitrou P, Romaguera D, Brown S, Bender A, et al. Operationalizing the 2018 World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) cancer prevention recommendations: a standardized scoring system. Nutrients. 2019;11:1572. https://doi.org/10.3390/nu11071572.
26. Sung JJY, Wong MCS, Lam TYT, Tsoi KKF, Chan VCW, Cheung W, et al. A modified colorectal screening score for prediction of advanced neoplasia: a prospective study of 5744 subjects. J Gastroenterol Hepatol. 2018;33:187–94. https://doi.org/10.1111/jgh.13835.
27. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94:496–509. https://doi.org/10.1080/01621459.1999.10474144.
28. Putter H, Spitoni C. Non-parametric estimation of transition probabilities in non-Markov multi-state models: the landmark Aalen-Johansen estimator. Stat Methods Med Res. 2018;27:2081–92. https://doi.org/10.1177/0962280216674497.
29. Main Data of the Seventh National Population Census. https://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817185.html. Accessed 4 Mar 2026. Cited 2026 Mar 4.
30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based

linkage analyses. Am J Hum Genet. 2007;81:559–75. https://doi.org/10.1086/519795.
31. R Core Team. R: A Language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2024. https://www.R-project.org/.
32. Briggs SEW, Law P, East JE, Wordsworth S, Dunlop M, Houlston R, et al. Integrating genome-wide polygenic risk scores and non-genetic risk to predict colorectal cancer diagnosis using UK Biobank data: population based cohort study. BMJ. 2022;e071707. https://doi.org/10.1136/bmj-2022-071707.
33. Thomas M, Sakoda LC, Hoffmeister M, Rosenthal EA, Lee JK, Van Duijnhoven FJB, et al. Genome-wide modeling of polygenic risk score in colorectal cancer risk. Am J Hum Genet. 2020;107:432–44. https://doi.org/10.1016/j.ajhg.2020.07.006.
34. Imperiale TF, Gruber RN, Stump TE, Emmett TW, Monahan PO. 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. https://doi.org/10.7326/M18-2390.
35. Naber SK, Kundu S, Kuntz KM, Dotson WD, Williams MS, Zauber AG, et al. Cost-effectiveness of risk-stratified colorectal cancer screening based on polygenic risk: current status and future potential. JNCI Cancer Spectr. 2020;4:pkz086. https://doi.org/10.1093/jncics/pkz086.
36. Cenin DR, Naber SK, De Weerdt AC, Jenkins MA, Preen DB, Ee HC, et al. Cost-effectiveness of personalized screening for colorectal cancer based on polygenic risk and family history. Cancer Epidemiol Biomark Prev. 2020;29:10–21. https://doi.org/10.1158/1055-9965.EPI-18-1123.
37. Xia C, Xu Y, Li H, He S, Chen W. Benefits and harms of polygenic risk scores in organised cancer screening programmes: a cost-effectiveness analysis. Lancet Reg Health West Pac. 2024;44:101012. https://doi.org/10.1016/j.lanwpc.2024.101012.

## Publisher's Note