Gastro Hep Advances 2024;3:151–161

ORIGINAL RESEARCH—CLINICAL

Colorectal Cancer Polygenic Risk Score Is Associated With Screening Colonoscopy Findings but Not Follow-Up Outcomes

Brian A. Sullivan,[1,2] Xuejun Qin,[1,2] Thomas S. Redding IV,[1] David Weiss,[3] Julie Upchurch,[1] Kellie J. Sims,[1] Jason A. Dominitz,[4,5] Anjanette Stone,[6] Belinda Ear,[1] Christina D. Williams,[1,2] David A. Lieberman,[7,8] and Elizabeth R. Hauser[1,2]

[1]Cooperative Studies Program Epidemiology Center-Durham, Durham VA Health Care System, Durham, North Carolina; [2]Division of Gastroenterology, Duke University, Durham, North Carolina; [3]Cooperative Studies Program Coordinating Center, Perry Point VA Medical Center, Perry Point, Maryland; [4]Division of Gastroenterology, VA Puget Sound Health Care System, Seattle, Washington; [5]Division of Gastroenterology, University of Washington School of Medicine, Seattle, Washington; [6]Cooperative Studies Program Pharmacogenomics Analysis Laboratory, Central Arkansas Veterans Health System, Little Rock, Arkansas; [7]Division of Gastroenterology, VA Portland Health Care System, Portland, Oregon; and [8]Division of Gastroenterology, Oregon Health & Science University, Portland, Oregon

BACKGROUND AND AIMS: Colorectal cancer (CRC) polygenic risk scores (PRS) may help personalize CRC prevention strategies. We investigated whether an existing PRS was associated with advanced neoplasia (AN) in a population undergoing screening and follow-up colonoscopy. **METHODS:** We evaluated 10-year outcomes in the Cooperative Studies Program #380 screening colonoscopy cohort, which includes a biorepository of selected individuals with baseline AN (defined as CRC or adenoma ≥10 mm or villous histology, or high-grade dysplasia) and matched individuals without AN. A PRS was constructed from 136 prespecified CRC-risk single nucleotide polymorphisms. Multivariate logistic regression was used to evaluate the PRS for associations with AN prevalence at baseline screening colonoscopy or incident AN in participants with at least one follow-up colonoscopy. **RESULTS:** The PRS was associated with AN risk at baseline screening colonoscopy ($P = .004$). Participants in the lowest PRS quintile had more than a 70% decreased risk of AN at baseline (odds ratio 0.29, 95% confidence interval 0.14–0.58; $P < .001$) compared to participants with a PRS in the middle quintile. Using a PRS cut-off of more than the first quintile to indicate need for colonoscopy as primary screening, the sensitivity for detecting AN at baseline is 91.8%. We did not observe a relationship between the PRS and incident AN during follow-up ($P = .28$). **CONCLUSION:** A PRS could identify individuals at low risk for prevalent AN. Ongoing work will determine whether this PRS can identify a subset of individuals at sufficiently low risk who could safely delay or be reassured about noninvasive screening. Otherwise, more research is needed to augment these genetic tools to predict incident AN during long-term follow-up.

Keywords: Colorectal Cancer; Polygenic Risk Score; Screening; Surveillance; Colonoscopy

Introduction

Colorectal cancer (CRC) is a preventable disease, yet remains the second leading cause of cancer death in the United States and Western Europe.[1,2] Several real-world limitations exist that reduce the effectiveness of current paradigms for CRC prevention. Most cases of CRC occur in those who are unscreened,[3,4] who may not be aware of their personal risk for CRC or have limited access due to colonoscopy-first approaches to screening (particularly in countries where colonoscopy is the predominant method for screening like the United States).[5] Furthermore, evidence suggests that many people have an aversion to undergoing screening through stool-based tests, which limits uptake and (long-term) adherence.[6,7] Finally, evidence supporting ongoing colonoscopy follow-up after a baseline examination is also limited, as most individuals will not develop CRC despite exposure to the costs and risks of repeated examinations.[8–12] Therefore, a clear need exists to improve CRC risk-stratification algorithms to better identify high-risk individuals who benefit from earlier and/or more intensive screening and surveillance, from those who are lower risk and may be appropriate for less intensive, less frequent, or noninvasive screening and surveillance.

Abbreviations used in this paper: **AN, advanced neoplasia; CI, confidence interval; CRC, colorectal cancer; CSP, Cooperative Studies Program; FIT, fecal immunochemical test; IQR, interquartile range; IRB, Institutional Review Board; LRA, low-risk adenomas; MAF, minor allele frequency; NSAID, nonsteroidal anti-inflammatory drug; OR, odds ratio; PRS, polygenic risk scores; SNPs, single nucleotide polymorphisms; VA, Veterans Affairs.**

Most current article

2772-5723
https://doi.org/10.1016/j.gastha.2023.10.001

An initial, noninvasive blood test has the potential to overcome these obstacles and enhance CRC screening rates. Although these biomarkers are currently in development,[13] another promising blood-based approach is to use genetic risk prediction tools to "triage" and inform Patient-Provider conversations around initial screening options and follow-up decisions.[14,15] A polygenic risk score (PRS) is a tool that can leverage fixed genetic risk factors to improve individualized CRC risk prediction in individuals without hereditary cancer syndromes. A PRS assumes that disease susceptibility is a function of a large set of genetic variants each with small effect (single nucleotide polymorphisms, "SNPs"), and summarizes the impact of the known variants on CRC risk that are present in an individual. Therefore, these scores may provide more accurate, personalized risk assessments. Prior case-control studies have shown that a PRS may help prioritize individuals for *first-time* CRC screening.[16–21] More recent studies in a German screening colonoscopy cohort have also shown for the first time that a PRS may provide improved estimates of CRC risk after a baseline examination, which could help tailor follow-up intervals.[22,23] However, since these studies mostly evaluated cancer outcomes, it remains unknown whether a PRS can characterize risk for important precancerous factors at baseline screening or during follow-up, or identify a relevant subset of individuals at sufficiently low risk who would be unlikely to benefit from a colonoscopy-based screening and follow-up paradigms.

Development of an effective genetic risk prediction tool has been limited by a lack of genomic databases with well-annotated, long-term screening colonoscopy and follow-up outcomes.[24] The Veterans Affairs (VA) Cooperative Studies Program (CSP) #380 is an independent prospective screening colonoscopy cohort with an associated biorepository that can uniquely address this limitation.[10,25] The present study aims to externally validate an existing PRS in CSP#380 by testing associations with prevalent or incident advanced neoplasia (AN) during baseline screening and follow-up, respectively. Improving CRC risk prediction could both reduce CRC risk and unnecessary healthcare utilization by more precisely tailoring CRC screening and follow-up resources.[26]

Methods

VA CSP#380 is an observational cohort study of long-term screening colonoscopy outcomes. Detailed methodology has been previously described.[10,25] In brief, this study enrolled 3121 Veterans aged 50–75 years from 13 VA Medical Centers between 1994 and 1997. Exclusion criteria were mostly selected to avoid a medical condition that could increase the risk or preclude benefit from CRC screening, including rectal bleeding, marked change in bowel habits, history of colonic disease (colitis, polyps, or cancer), or a colonic examination within the previous 10 years. Participants completed detailed questionnaires regarding medical history, including family history of cancer, and then underwent screening colonoscopy by study investigators. Importantly, the quality of baseline screening examinations was documented and well within current quality metric guidelines.[25] A longitudinal study was then conducted to assess 10-year clinical outcomes of those undergoing screening and programmatic follow-up.[10] Documentation was obtained from each study site regarding the findings of each colonoscopy, including polyp number, size, and histology. Validation occurred by review of colonoscopy and pathology records ascertained through VA electronic medical records, which includes collection of reports for procedures performed outside of the VA when available.

Biorepository and Genotyping

As part of the initial CSP#380 study, blood samples were obtained from 815 selected Veterans and stored in a biorepository.[27] A nested case-control design was used for efficiency purposes, and so Veterans were allocated to the biorepository based on the presence of CRC or adenoma size $\geq$10 mm at baseline colonoscopy (n = 226) together with age-matched and sex-matched controls (n = 589). Genotyping was completed on extracted DNA from lymphocytes (ie, blood) using the Illumina Infinium Omni2.5-8 v1.3 Beadchip. Genetic ancestry, an important potential confounder in genetic studies, was accounted for by including 2 genetic principal components. Principal Components Analysis was performed using Hapmap3 with linkage disequilibrium pruning at R^2 >0.05 (Plink). Principal Components Analysis was performed on 24,006 markers with minor allele frequency >0.4 using the R package flashpcaR. European ancestry was defined by genetic clustering with Europeans from HapMap.

DNA quantity and/or quality was not sufficient for 203 samples, so 612 (European ancestry: 500 participants; non-European ancestry: 112 participants) remained in the CSP#380 genetic analysis dataset. Genotyping showed excellent quality of markers for the analytic dataset participants: 99% markers with < 5% missing data. Of these 612 individuals, 2 participants were excluded due to relatedness. Baseline characteristics of the 610 included compared to excluded participants are reported in Table A1 to acknowledge potential bias due to incomplete clinical or genomic data. This protocol is approved by the Durham VA Institutional Review Board under the CSP#380b study (#01797/0002).

Construction of the Polygenic Risk Score

A PRS prespecified from the existing literature was calculated for each CSP#380 participant in the analytic dataset by established methods.[17–23,28] In brief, this PRS is based on 140 known SNPs, and their literature effect sizes, that have been shown to be independently associated with CRC based on studies in large genetic consortia through collaborative meta-analyses adjusted for age, sex, study, and principal components.[28] SNPs not included on our genotyping platform were imputed based on robust validated methodology to estimate and account for missing data.[29] Specifically, genotype imputation was performed using a pipeline developed by the Million Veteran Program using the TOPMed reference panel.[30,31] As a measure of imputation quality and reliability, the imputation R^2 was 0.988 in the 628,849 directly genotyped markers. After

analysis using the PRSice software, of the 140 SNPs in the original PRS, 136 were available for this study and used in the calculation of the PRS values (regardless of minor allele frequency). A specific PRS was assigned to each individual based on the total number of risk alleles present per CRC-risk SNP, with a range per SNP between 0 and 2 points (0 if the risk allele is not present, 1 if heterozygous, and 2 if homozygous). Each SNP was weighted based on its published effect size for CRC by multiplying the total points per SNP by the published log (odds ratio [OR]). The final score was the sum total number of points across all CRC-risk SNPs present in each individual.

Prevalent Advanced Neoplasia at Baseline Screening Colonoscopy

Our primary outcome is AN (defined as invasive CRC or an adenoma ≥10 mm, or an adenoma with villous histology or high-grade dysplasia). Cases of prevalent AN were defined as AN detected at baseline screening colonoscopy, or during colonoscopy within 6 months of the baseline colonoscopy as per the initial study protocol (Figure 1). Cases of AN were evaluated against baseline controls regardless of whether participants had follow-up or not. For analyses of prevalent AN, controls were defined as participants with no AN at the baseline colonoscopy, and if the participant had follow-up, no AN throughout the entire 10-year observation period. For this analysis, participants with only AN at follow-up were excluded, as they would not fit the definition of case or control. Thus, of 610 biorepository participants, there were 171 cases of prevalent AN and 423 controls without AN ever during the observation period, for a total of 594 included participants.

Incident Advanced Neoplasia at Follow-up Colonoscopy

The analysis of incident AN was restricted to only participants with at least one follow-up colonoscopy after baseline to allow for potential AN detection during the observation period (Figure 1). Cases of incident AN were defined as any instance of AN detected on any follow-up colonoscopy at least 6 months after the initial examination (regardless of baseline findings). Controls were defined as participants with no AN on any examination across the entire 10-year observation period. For

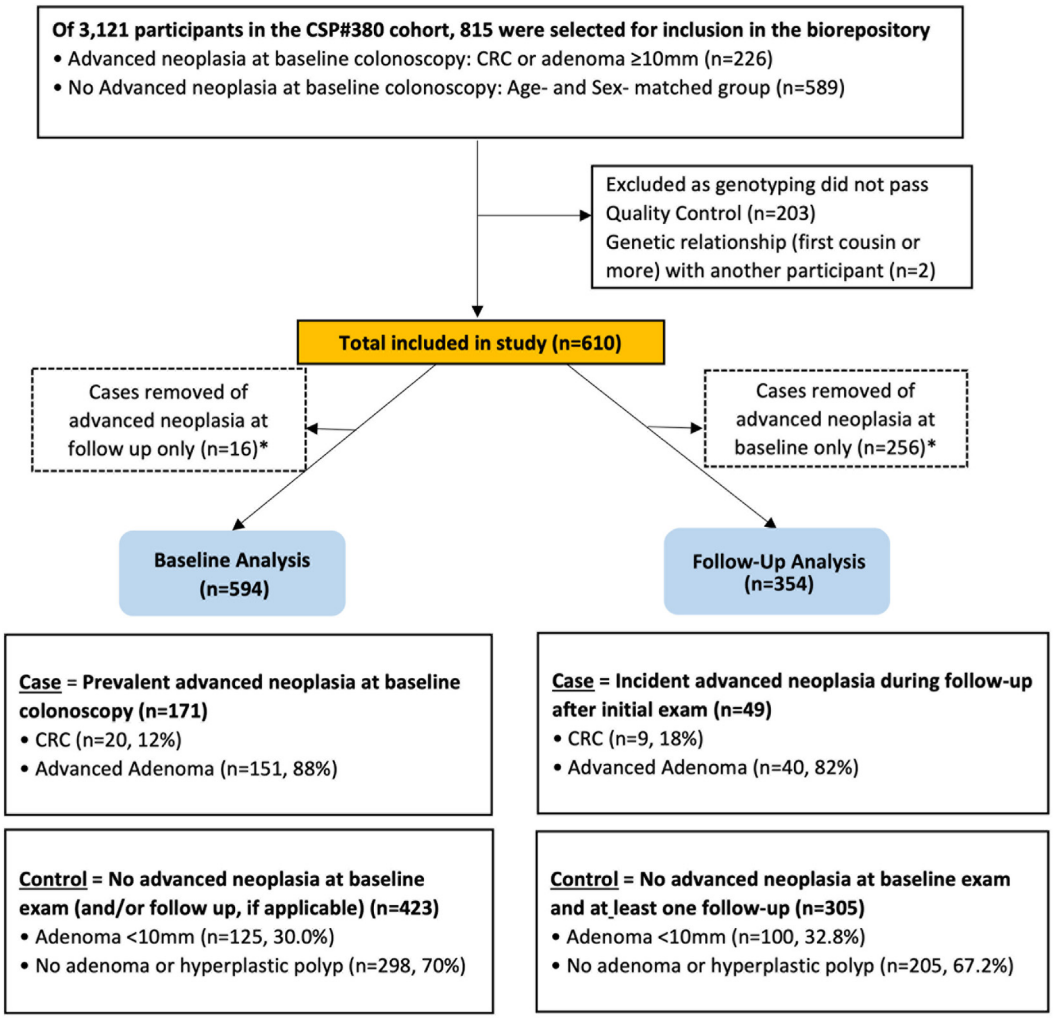

Figure 1. CONSORT flow diagram: Biorepository allocation from CSP#380 participants (n = 815).

this analysis, participants with only AN at baseline were excluded, as they would not fit the definition of case or control. Thus, of 610 biorepository participants, there were a total of 354 participants with any follow-up included, including 49 cases of incident AN and 305 controls without AN ever during the observation period.

Statistical Analysis

We report descriptive statistics and associations of the PRS with AN as a continuous variable in the overall included cohort, as well as in stratified analyses, at a statistical significance threshold of $P < .05$. Additionally, although prior studies using this PRS evaluated outcomes in participants at each decile of the score,[28] due to power considerations, we elected *a priori* to evaluate the PRS by quintiles to determine risk groups that the PRS may have the most utility. Weighted PRS quintile cut-offs were separately determined for each outcome, prevalent and incident AN, based on the range of the PRS (ie, the distribution of weighted risk alleles) across all participants included in the respective control group.[19–23,28] Given that 4 comparisons are tested against a reference level between the 5 PRS categories (quintiles), a Bonferroni multiple-comparison adjustment was applied for a nominal P value of .0125 to declare statistical significance.

For the analysis of prevalent or incident AN, we fit a mixed model to estimate the association between the PRS and AN to account for the longitudinal nature of our study, which includes participants undergoing various intensity of follow-up.[10] This statistical model includes all colonoscopies for each person over the 10-year follow-up period and accounts for prior endoscopic findings and varying time under surveillance. We also controlled for age at baseline (for prevalent AN analysis) and age at last colonoscopy or first colonoscopy with AN (for the incident AN analysis) and sex. Subgroup analyses were performed by stratifying individuals by their age, sex, ancestry, and baseline colonoscopy findings (AN vs lower risk findings). Due to smaller sample sizes, these stratified analyses are considered exploratory. Finally, given that the design of the initial biorepository over-represented participants with a family history of CRC, sensitivity analyses were performed excluding participants with a family history of CRC to more closely approximate an "average risk" screening population.

Based initial reviews of the PRS distribution between cases and controls, which suggested that polygenic risk-based discrimination may be most effective at the fringes of these scores and not across the entire spectrum of risk, "Area Under the Curve" of the "Receiver Operating Characteristic" curve analyses were not conducted. To ascertain whether the PRS may select a subset of individuals in whom baseline screening colonoscopy could be safely deferred, sensitivity and specificity were calculated for AN and CRC using cut-points at each PRS quintile for the prevalent AN analysis only (where results above the respective cut-point are "positive" and results below the cut-point are "negative"). Test characteristics were calculated at each cut-point, where "true positives" and "false negatives" were defined as cases of AN above and below the respective threshold, respectively, while "true negatives" and "false positives" were defined as controls below and above the respective cut-point, respectively. These results were then compared with those that would have been obtained using colonoscopy alone as per the CSP#380 protocol to calculate number of colonoscopies saved at each cut-point (ie, those who would be screened by another strategy initially if the PRS was "negative" as determined by being below the respective quintile threshold). Given that estimates of positive predictive value and negative predictive value depend on the prevalence of disease in the tested population, which is inherently distorted in a case-control selection design, we elected not to include these calculations in this analysis.

Results

As expected based upon the initial recruitment strategy for the biorepository, the 610 participants with genotyped blood samples (19.5% of the overall 3121 original cohort) were older (64.1 vs 62.6, $P < .001$), more likely to have a first degree family history of CRC (22.1% vs 11.9%, $P < .001$), and had a higher prevalence of AN (including CRC) on baseline screening colonoscopy (28.0% vs 6.3%, $P < .001$), compared to those not included in the CSP#380 biorepository (Table A1). There were no significant differences in representation based on sex or self-reported race. Of those included in each analysis, Table 1 describes their baseline demographic and clinical characteristics.

Prevalent AN Analysis

As described above and in Figure 1, in the analysis of prevalent AN on the baseline screening examination, there were 171 cases of AN at baseline screening colonoscopy and 423 controls without AN ever during the 10-year observation period (regardless of follow-up). Cases were more likely to be older ($P = .05$), current smokers ($P = .04$), or have a family member with CRC ($P = .08$) when compared with controls (Table 1). The median PRS (interquartile range) was 3.68 (0.22) for cases with AN at baseline screening colonoscopy and 3.64 (0.25) for control participants with no AN ever on at least a baseline examination. Additionally, the median number of CRC-risk SNP risk alleles was 127 (8.5) and 125 (11), respectively.

When evaluating the PRS as a categorical variable, there were no baseline participant characteristics significantly different across PRS quintiles (Table A2). However, some differences did approach statistical significance. For example, there were differences in the distribution of self-reported race and ethnicity ($P = .08$) across PRS quintiles, with White (non-Hispanic) participants seemingly less likely to be in the lowest quintile. Furthermore, those with a family history of CRC were more likely to be in the highest PRS quintile ($P = .05$).

As shown in Table 2, the PRS as a continuous variable was significantly associated with AN risk at baseline screening colonoscopy ($P = .004$). While a one-point increase in the PRS was associated with an OR of 5.01 (95% confidence interval [CI], 1.69–15.18), the full range of the PRS is only approximately one point (3.2–4.2). In stratified analyses intended to explore whether the PRS as a continuous variable was most impactful in certain populations

Table 1. Selected Characteristics of the Study Population, Stratified by Timing of Advanced Neoplasia

Variable	Baseline controls N = 423	Prevalent advanced neoplasia cases N = 171	*P* value[a]	Follow-up controls N = 305	Incident advanced neoplasia cases N = 49	*P* value[a]
Age, mean (SD)	71.63 (7.77)	73 (6.7)	.045	76.42 (7.77)	79.11 (5.76)	.02
Male sex, N (%)	409 (96.69)	168 (98.25)	.45	295 (96.72)	48 (97.96)	.98
Self-reported race and ethnicity, N (%)			.21			.16
White, non-Hispanic	341 (80.81)	143 (83.63)		249 (81.91)	42 (87.5)	
Black, non-Hispanic	46 (10.9)	11 (6.43)		34 (11.18)	1 (2.08)	
Hispanic	23 (5.45)	8 (4.68)		15 (4.93)	2 (4.17)	
American Indian/Alaskan Native	9 (2.13)	5 (2.92)		4 (1.32)	2 (4.17)	
Asian	3 (0.71)	4 (2.34)		2 (0.66)	1 (2.08)	
Education, N (%)			.26			.50
Under 7 y schooling	13 (3.07)	3 (1.75)		7 (2.3)	1 (2.04)	
Junior high school	37 (8.75)	18 (10.53)		27 (8.85)	4 (8.16)	
Some high school	41 (9.69)	23 (13.45)		27 (8.85)	8 (16.33)	
High school graduate	99 (23.4)	50 (29.24)		71 (23.28)	15 (30.61)	
Some college	139 (32.86)	44 (25.73)		105 (34.43)	13 (26.53)	
College graduate	65 (15.37)	20 (11.7)		44 (14.43)	4 (8.16)	
Completed graduate training	29 (6.86)	13 (7.6)		24 (7.87)	4 (8.16)	
Current smoker, N (%)	82 (19.39)	47 (27.49)	.04	50 (16.39)	13 (26.53)	.13
Past smoker, N (%)	235 (68.91)	85 (68)	.94	177 (69.41)	30 (83.33)	.13
Family history, N (%)	104 (24.59)	30 (17.54)	.08	82 (26.89)	5 (10.2)	.02
Weighted PRS, median (IQR)	3.64 (0.25)	3.68 (0.22)	.003	3.64 (0.25)	3.67 (0.23)	.23
Unweighted PRS, median (IQR)	125 (11)	127 (8.5)	.005	125 (10)	128 (6)	.23

IQR, interquartile range; N, total number; PRS, Polygenic Risk Score; SD, standard deviation.
[a]Compared to respective control group.

(Table 2), the impact on baseline AN risk of the weighted PRS was most pronounced in those with a family history of CRC ($P = .04$), European ancestry ($P = .001$), and participants aged 65–75+ years ($P = .01$). The PRS was not significantly associated with baseline AN risk in the non-European or age 50–64 subgroups.

Figure 2A demonstrates the overall distribution of the PRS between participants with AN (cases) and without (controls) at baseline screening colonoscopy, which suggests more meaningful differentiation in scores at the lowest PRS quintile. Specifically, there was a smaller proportion of cases with baseline AN in the lowest PRS quintile (8.2%), compared to those without AN at baseline (20.0%), thus cases were more likely to be distributed across higher PRS scores. As highlighted in Table 3, participants with a PRS in the lowest quintile had more than a 70% decreased risk of AN at baseline screening colonoscopy (OR 0.29, 95% CI 0.14–0.58; $P < .001$) compared to participants with a PRS in the middle quintile, which met criteria for multiple-comparison adjusted statistical significance ($P < .0125$). Only one CRC was identified in the lowest PRS quintile in an individual aged 65 years, compared to 19 cancers in the other 4 PRS quintiles (with 4, 7, 1, and 7 cancers in those quintiles from lowest to highest).

In sensitivity analyses evaluating the PRS as a continuous variable on prevalent AN risk after excluding those with a family history of CRC (Table A3), results were largely the same as the full cohort ($P = .02$). Similarly, the associations remained between the PRS and prevalent AN risk at baseline in those of European ancestry ($P = .007$) and participants aged 65–75+ years ($P = .05$) in the stratified analyses. As shown in Table A4, in those without a family history, participants with a PRS score in the lowest quintile had more than a 75% decreased risk of AN at baseline colonoscopy (OR 0.24, 95% CI 0.11–0.50, $P < .001$), compared to those in the middle PRS quintile, meeting criteria for multiple-comparison adjusted significance. While participants in the fourth PRS quintile seemingly had a lower risk of baseline AN (OR 0.50, 95% CI 0.26–0.94, $P = .03$), the P value was not below the nominal significance threshold of .0125.

Using the PRS as a "triage" test to indicate a "positive" result if the individual's score is more than the first quintile cut-off, where an individual would be recommended to undergo colonoscopy as the initial screening modality, the sensitivity and specificity for AN detection is 91.8% and 20.1%, respectively, and for CRC is 95.0% and 17.3%, respectively (Table 4). Using this threshold, there would be a reduction of 16.6% in the use of colonoscopy as the initial screening modality (ie, 99 of 594 participants with a "negative" PRS score could undergo delayed or alternative, noncolonoscopic screening). Additional test characteristics calculated at each quintile cut-point are shown in Table 4.

Table 2. Association Between PRS (Continuous Variable) and Prevalent or Incident Advanced Neoplasia in the Overall CSP#380 Cohort, and Stratified by Family History, Ancestry, Age, or Baseline Examination Findings

	Number of participants without advanced neoplasia N (%)	Number of participants with advanced neoplasia N (%)	Advanced neoplasia odds ratio (95% CI)[a]	*P* value
Prevalent advanced neoplasia outcomes, overall and stratified by risk group				
Overall cohort (n = 594)	423 (71.21%)	171 (28.79%)	5.01 (1.69, 15.18)	.004
Family history of CRC in first degree relative (n = 134)	104 (77.6%)	30 (22.4%)	12.6 (1.20, 155)	.04
Ancestry				
European (n = 484)	341 (70.5%)	143 (29.5%)	7.80 (2.30, 27.52)	.001
Non-European (n = 110)	82 (74.5%)	28 (25.5%)	0.81 (0.06, 10.21)	.87
Age				
Age 50–64 (n = 99)	80 (80.8%)	19 (19.2%)	1.29 (0.48, 435.4)	.14
Age 65–75+ (n = 492)	342 (69.5%)	150 (30.5%)	4.44 (1.39, 14.52)	.01
Incident advanced neoplasia outcomes, overall and stratified by risk group				
Overall cohort (n = 354)	305 (86.2%)	49 (13.8%)	2.64 (0.45, 15.82)	.28
Family history of CRC in first degree relative (n = 87)	82 (84.3%)	5 (15.7%)	11.9 (0.02, >1000)	.47
Ancestry				
European (n = 291)	249 (85.6%)	42 (14.4%)	1.80 (0.28, 15.4)	.58
Non-European (n = 63)	56 (88.9%)	7 (11.1%)	0.08 (0.0003, 10.0)	.34
Age				
Age 50–64 (n = 24)	24 (100%)	0 (0%)	0.37 (0.01, 11.1)	.58
Age 65–75+ (n = 329)	280 (85.1%)	49 (14.9%)	1.42 (0.21, 15.2)	.74
Baseline colonoscopy findings				
No adenomas/low-risk adenomas (n = 296)[b]	284 (95.9%)	12 (4.1%)	0.78 (0.004, 30.3)	.91
3+ nonadvanced adenomas(n = 25)	21 (84%)	4 (16%)	0.69 (<0.001, >1000)	.96
Advanced neoplasia (n = 33)	0 (0%)	33 (100%)	0.62 (0.07, 6.13)	.66

CRC, colorectal cancer; PRS, Polygenic Risk Score.
[a]Models adjusted for ancestry (based on genetic ancestry by principal component analysis), sex, and age [at last colonoscopy or first colonoscopy with advanced neoplasia].
[b]Defined as no adenomas or 1–2 small (<10 mm) adenomas.

Incident AN Analysis

In the analysis of incident AN, as described above, there were 49 cases of incident AN (with or without baseline AN) on a follow-up colonoscopy and 305 controls without AN at baseline nor at follow-up (Figure 1). Older age ($P = .02$) and a first degree family history of CRC ($P = .02$) were more common in cases than in controls (Table 1). The median (interquartile range) PRS was 3.67 (0.23) in cases with AN on at least one follow-up examination and 3.64 (0.25) in control participants with no AN at baseline nor at follow-up. The median number of CRC-risk SNP risk alleles was 128 (6) and 125 (10) in these participants with and without incident AN, respectively.

The PRS as a continuous variable was not significantly associated with AN risk at follow-up colonoscopy ($P = .28$) (Table 2). The PRS was also not significantly associated with incident AN during follow-up in stratified analyses by family history, age, ancestry, or baseline colonoscopy findings. As shown in Figure 2B, a smaller proportion of cases with incident AN on a follow-up examination were in the lowest PRS quintile (12.2%), compared to controls without AN ever during follow-up (20.0%), suggesting more cases were represented in higher PRS quintiles. Nevertheless, we did not observe a significant multiple-comparison adjusted relationship between the PRS and incident AN during follow-up (Table 3), including in the lowest PRS quintile (OR 0.77, 95% CI 0.38–1.50; $P = .46$). The number of baseline CRC cases in each PRS quintile from lowest to highest was 1, 3, 3, 1, and 1, respectively.

Finally, in sensitivity analyses including only those without a family history of CRC yielded similar results (Tables A3 and A4). The PRS was also not associated with incident AN during follow-up in these participants ($P = .34$) nor in any stratified analyses.

Discussion

We sought to externally evaluate the discriminatory performance of an existing CRC-risk PRS for estimating risk of AN, including CRC, in the CSP #380 Veteran screening colonoscopy population.[19,28,32] Findings from this nested case-control study provide further support for use of a PRS as a risk stratification tool in men undergoing initial CRC screening, particularly at the lowest end of the genetic risk

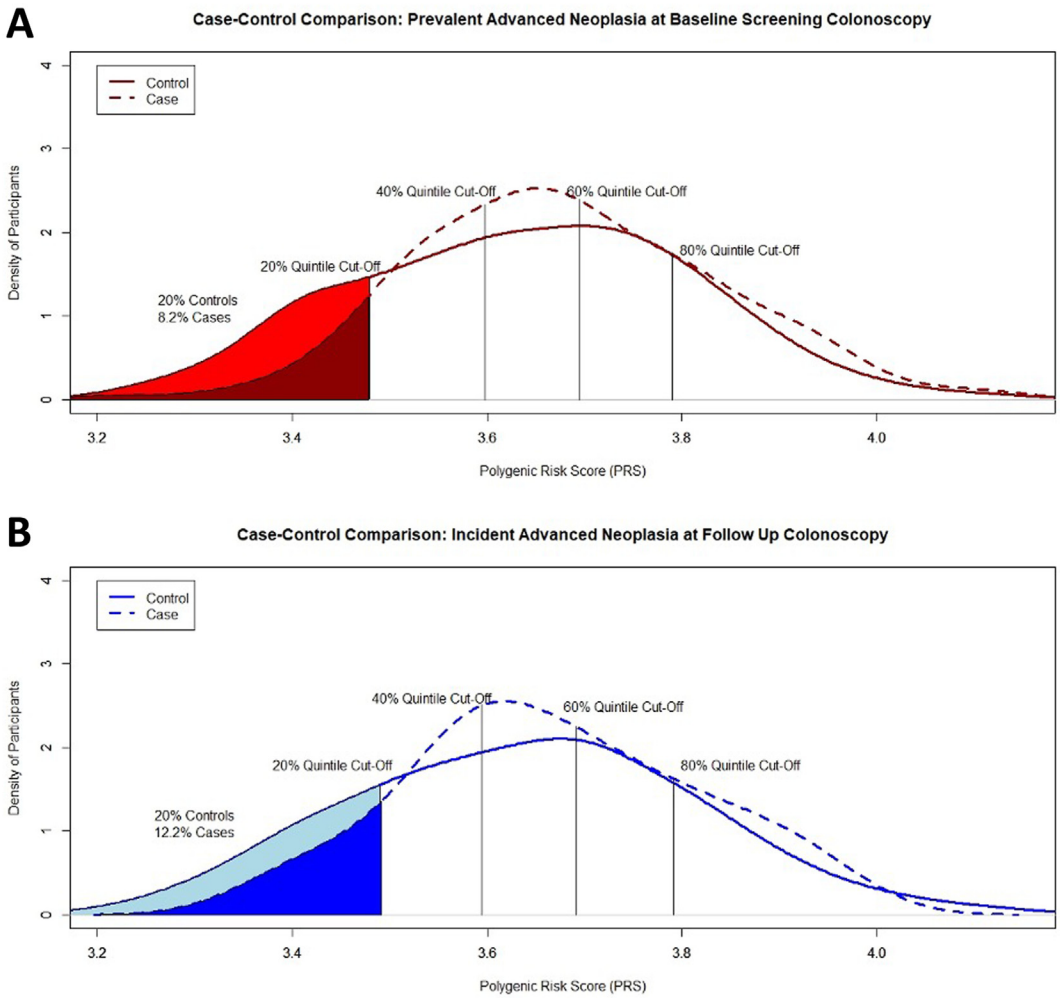

Figure 2. (A) PRS distribution between participants with and without prevalent advanced neoplasia. (B) PRS distribution between participants with and without incident advanced neoplasia. The x-axis indicates the distribution of the polygenic risk score across all cases (dashed lines) and controls (solid lines). The y-axis indicates the density of participants. Solid areas and corresponding percentages are the proportion of cases and controls with a PRS below 20% in the controls. PRS, Polygenic Risk Score.

spectrum. However, the PRS was not significantly associated with AN risk during follow-up after the baseline screening exam.

We found that an existing PRS had discriminatory value for CRC screening in an independent male veteran population. Participants with a PRS value in the lowest quintile had more than a 70% decreased risk of AN at baseline screening colonoscopy (OR 0.29, 95% CI 0.14–0.58; $P < .001$) compared to those with a PRS in the middle quintile. Applied in clinical practice as a risk stratification tool, these "low-risk" individuals may not require invasive testing with colonoscopy as the initial screening modality or may be able to defer initiation of screening until a later age than is recommended for the general population. Based on our study, 17% of colonoscopies could be deferred in individuals with a PRS below a cut-off "triage" value of less than the first quintile. There would be few missed lesions since this cut-off would detect almost 92% of AN and 95% of CRC. The one CRC case in the lowest PRS quintile occurred in a 65-year-old Veteran, which conceivably may have been detected or prevented even with delayed screening initiation after age 45. Indeed, it is possible that fecal immunochemical test or other noninvasive tests would be complementary if administered to the group deemed low-risk based on PRS.[14,33,34] Primary Care Providers may be encouraged to more commonly recommend a fecal immunochemical test–first strategy for a subset of their patients, based on patient preferences and knowledge of low genetic risk.[35] This efficient use of screening resources could then help ensure uptake of screening in higher risk individuals by increasing capacity and improving access.[14] Overall, our results suggest that a PRS-based screening strategy has utility for the detection of AN, but more work is needed to achieve a performance that justifies use in clinical practice.

Prior studies in cohorts of European ancestry also found that individuals with the lowest PRS had low risk for AN and

Table 3. Association Between PRS (Categorical Variable) and Prevalent or Incident Advanced Neoplasia

	Number of participants without advanced neoplasia[a] N (%)	Number of participants with advanced neoplasia N (%)	Baseline advanced neoplasia odds ratio (95% CI)[b]	*P* value
Prevalent advanced neoplasia outcomes (n = 594)				
Weighted PRS by quintile				
First quintile (n = 99)	85 (85.9%)	14 (14.1%)	0.29 (0.14, 0.58)	<.001
Second quintile (n = 121)	84 (69.4%)	37 (30.6%)	0.86 (0.50, 1.47)	.58
Third quintile (n = 130)	85 (65.4%)	45 (34.6%)	Ref	ref
Fourth quintile (n = 117)	84 (71.8%)	33 (28.2%)	0.75 (0.43, 1.31)	.32
Fifth quintile (n = 127)	85 (66.9%)	42 (33.1%)	0.95 (0.56, 1.61)	.85
Incident advanced neoplasia outcomes (n = 354)				
Weighted PRS by quintile				
First quintile (n = 67)	61 (91.0%)	6 (9.0%)	0.77 (0.38, 1.5)	.46
Second quintile (n = 73)	61 (83.6%)	12 (16.4%)	0.86 (0.46, 1.59)	.62
Third quintile (n = 73)	61 (83.6%)	12 (16.4%)	Ref	ref
Fourth quintile (n = 70)	61 (87.1%)	9 (12.9%)	0.99 (0.52, 1.88)	.99
Fifth quintile (n = 71)	61 (85.9%)	10 (14.1%)	0.76 (0.40, 1.38)	.37

PRS, Polygenic Risk Score.
[a]Quintile cut-offs calculated by the distribution of the PRS, among healthy control participants, thus equivalent numbers of controls are represented in each quintile.[19–23,28]
[b]Models adjusted for ancestry (based on genetic ancestry by principal component analysis), sex, and age [at last colonoscopy or first colonoscopy with advanced neoplasia].

CRC.[28] In a study using the largest available genetic databases, a PRS categorized a similar proportion of CRC cases (8.1%) in the bottom 20% of genetic risk as in our study (8.2%).[28] Our findings are also consistent with 2 studies of screening colonoscopy populations from Germany. We found an OR of 0.29 for prevalent AN risk in the lowest PRS *quintile* (compared to the middle quintile), comparable to ORs for AN and CRC risk, respectively, of approximately 0.37 and 0.45 in the lowest PRS *tertile* (compared to those in the highest tertile).[17,21] Weigl et al concluded that individuals with the lowest PRS could delay screening by as much as 10 years given the delay in attaining similar risk for AN as those in other PRS risk categories. Jeon et al[20] reported similar findings using a CRC PRS within a large genetic consortium, which translated to significant potential for delayed screening initiation in analyses based on age-adjusted CRC risk. Prospective studies are needed in diverse populations of non-European ancestry to evaluate the effectiveness of genetic risk assessment for early CRC detection and prevention.[36–38]

Certain population subgroups may benefit most from personalized risk assessments. Increasing age and family history add to the discriminatory performance of the PRS in our study and others.[20,21,28,37,39] Low penetrant genetic factors may increase CRC risk as people age into their seventh and eighth decades, accounting for the age effect. A

Table 4. Performance of the PRS as a Tool for Triaging Primary Screening Colonoscopy[a]

	Advanced neoplasia at screening (n = 171)		CRC at screening (n = 20)		Number of colonoscopies (n = 594)
PRS test "positivity" cut-point of CRC risk[b]	Sensitivity	Specificity	Sensitivity	Specificity	Colonoscopy saving as initial test[c]
Quintile 1	91.8%	20.1%	95.0%	17.3%	99/594 = 16.6%
Quintile 2	70.2%	40.0%	75.0%	37.8%	220/594 = 37.0%
Quintile 3	43.9%	60.0%	40.0%	58.9%	350/594 = 58.9%
Quintile 4	24.6%	79.9%	35.0%	78.9%	467/594 = 78.6%

CRC, colorectal cancer; PRS, Polygenic Risk Score.
[a]Using a cut-off PRS, of greater than the respective quintile to indicate a "positive" triage test where an individual should undergo colonoscopy for primary screening.
[b]Colonoscopy performed in individuals with a PRS, above the respective quintiles would have the following sensitivity and specific to detect AN and CRC.
[c]Reduction in colonoscopy demand as the initial test if individuals with a "negative" PRS, below the respective quintile undergo alternative screening strategies.

PRS may also be an ideal surrogate to better assess heritable risk when family history is unavailable, especially since family history is often difficult to obtain in routine practice and potentially becoming increasingly more inaccurate as widespread screening reduces CRC incidence in family members. Indeed, genetic risk factors together with other factors (smoking, nonsteroidal anti-inflammatory drug use, obesity) could contribute to an overall more robust risk assessment.[16,17,20,32,40] Ongoing collaborative studies in diverse populations (including non-European ancestries) seek to create a more accurate, generalizable CRC risk prediction tool for early CRC detection and prevention that is based on an optimal cluster of genetic variants together with relevant demographic and clinical factors.

We hypothesized that PRS might also be discriminatory for predicting which individuals might be at risk for AN or CRC during follow-up after an initial examination. Our study is uniquely able to evaluate this question because individuals with baseline colonoscopy were followed closely over the next 10 years with well-annotated surveillance colonoscopy findings, including advanced precancerous and CRC outcomes. Prior work by Guo et al[23] found that 10 years after a negative screening colonoscopy, adjusted ORs for CRC were 0.44 (95% CI, 0.29–0.68) in individuals in the lowest tertile of genetic risk compared to unscreened individuals. The authors concluded that follow-up intervals after a negative screening colonoscopy could potentially be prolonged even further for people with a low PRS. In the same population, these authors found that individuals with low-risk adenomas at baseline and a PRS in the lowest tertile were at sufficiently low risk for CRC up to 10 years, but if these individuals with low-risk adenomas had medium or high tertiles of genetic risk, then this risk reduction after colonoscopy lasted only through ~6 years.[22] However, our results failed to demonstrate a similar association between the PRS and risk for incident AN/CRC during surveillance. There are several possible explanations. The impact of intensive surveillance colonoscopy may have reduced the overall risk for AN, by detecting and removing small polyps before AN could develop. Our sample size is relatively small, and in particular, incident CRCs were uncommon and may have been prevented by surveillance. Additional studies are needed to determine if genetic risk assessment could be augmented with additional markers from blood and/or colonic tissues to improve risk-stratification during surveillance. In addition, CRC risk prediction tools will need to account for the variable quality of colonoscopy.[41,42]

Our study has important limitations. Our cohort has a small sample size and is mostly of European ancestry, male, and composed of mainly older participants, which limits generalizability. Furthermore, this study is a nested case-control design, which may lead to selection bias due to a higher prevalence of AN compared to a general population. Although our study demonstrates that a low score on the PRS may indicate *lower* risk individuals, whether these individuals are truly *low-risk* remains unknown. More research is needed to clarify appropriate population-based PRS standards for reference during risk comparisons prior to implementation of any PRS-based screening strategies. Finally, this study was performed in an era where recognition and evaluations for hereditary CRC syndromes was not widespread, and many participants have since died, thus precluding these genetic evaluations. On the other hand, strength of the CSP#380 repository is that it was formed based on a prospective, well-defined protocol and with well-annotated clinical data at baseline and follow-up. Therefore, this unique dataset was created to minimize potential bias from these study designs, such as by leveraging a control group that did not manifest AN over 10 years, and distinguishes this cohort from other cross-sectional studies which are only able to study a single time point during screening. And although the small sample size precluded risk-adjusted estimates for ages to start screening or undergo follow-up based on genetic findings, together with other studies, our results demonstrate feasibility and biologic plausibility for developing PRS-based strategies for individualized CRC screening and follow-up. This discovery work will inform studies in larger databases that seek to develop more comprehensive risk scores based on clinical and genetic factors that are applicable across broader racial and ethnic backgrounds.[38]

In conclusion, we show in an independent population that individuals in the lowest quintile of a known PRS had significantly lower risk for AN at screening colonoscopy compared to those with higher PRS scores. Ongoing work is needed to confirm whether this PRS can identify a subset of individuals at sufficiently low risk across diverse populations who could safely delay or undergo less frequent or noninvasive screening. On the other hand, we observed a lack of discriminatory performance for the PRS in individuals undergoing follow-up after a baseline screening colonoscopy. More research is needed in diverse populations undergoing screening to augment current blood-based risk prediction tools, potentially with genetic information from colonic tissues, to develop cost-effective individualized follow-up strategies.[43] Ultimately, a validated risk prediction tool is needed which demonstrates applicability across diverse demographic groups and integrates into electronic health records. This tool could provide real-time personalized recommendations for CRC prevention as individuals age or acquire medical comorbidities.

Supplementary Materials

Material associated with this article can be found in the online version at https://doi.org/10.1016/j.gastha.2023.10.001.

References

1. Siegel RL, Wagle NS, Cercek A, et al. Colorectal cancer statistics, 2023. CA Cancer J Clin 2023;73:233–254.
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–249.
3. Doubeni CA, Fedewa SA, Levin TR, et al. Modifiable failures in the colorectal cancer screening process and their association with risk of death. Gastroenterology 2019;156:63–74.e6.
4. Meester RG, Doubeni CA, Lansdorp-Vogelaar I, et al. Colorectal cancer deaths attributable to nonuse of screening in the United States. Ann Epidemiol 2015; 25:208–213.e1.
5. May FP, Yang L, Corona E, et al. Disparities in colorectal cancer screening in the United States before and after implementation of the Affordable Care Act. Clin Gastroenterol Hepatol 2020;18:1796–1804.e2.
6. Jones RM, Woolf SH, Cunningham TD, et al. The relative importance of patient-reported barriers to colorectal cancer screening. Am J Prev Med 2010;38:499–507.
7. Cusumano VT, May FP. Making FIT count: maximizing appropriate use of the fecal immunochemical test for colorectal cancer screening programs. J Gen Intern Med 2020;35:1870–1874.
8. Gupta S, Lieberman D, Anderson JC, et al. Recommendations for follow-up after colonoscopy and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastrointest Endosc 2020;91:463–485.e5.
9. Lieberman D, Gupta S. Does colon polyp surveillance improve patient outcomes? Gastroenterology 2020; 158:436–440.
10. Lieberman D, Sullivan BA, Hauser ER, et al. Baseline colonoscopy findings associated with 10-year outcomes in a screening cohort undergoing colonoscopy surveillance. Gastroenterology 2020;158:862–874.e8.
11. Kobe EA, Sullivan BA, Qin X, et al. Longitudinal assessment of colonoscopy adverse events in the prospective Cooperative Studies Program no. 380 colorectal cancer screening and surveillance cohort. Gastrointest Endosc 2022;96:553–562.e3.
12. Sullivan BA, Qin X, Miller C, et al. Screening colonoscopy findings are associated with noncolorectal cancer mortality. Clin Transl Gastroenterol 2022;13:e00479.
13. Shaukat A, Levin TR. Current and future colorectal cancer screening strategies. Nat Rev Gastroenterol Hepatol 2022;19:521–531.
14. Chen H, Shi J, Lu M, et al. Comparison of colonoscopy, fecal immunochemical test, and risk-adapted approach in a colorectal cancer screening trial (TARGET-C). Clin Gastroenterol Hepatol 2023;21:808–818.
15. Chiu HM, Ching JY, Wu KC, et al. A risk-scoring system combined with a fecal immunochemical test is effective in screening high-risk subjects for early colonoscopy to detect advanced colorectal neoplasms. Gastroenterology 2016;150:617–625.e3.
16. Archambault AN, Su YR, Jeon J, et al. Cumulative burden of colorectal cancer-associated genetic variants is more strongly associated with early-onset vs late-onset cancer. Gastroenterology 2020;158:1274–1286.e12.
17. Carr PR, Weigl K, Edelmann D, et al. Estimation of absolute risk of colorectal cancer based on healthy lifestyle, genetic risk, and colonoscopy status in a population-based study. Gastroenterology 2020;159:129–138.e9.
18. Hsu L, Jeon J, Brenner H, et al. A model to determine colorectal cancer risk using common genetic susceptibility loci. Gastroenterology 2015;148:1330–1339.e14.
19. Huyghe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet 2019;51:76–87.
20. Jeon J, Du M, Schoen RE, et al. Determining risk of colorectal cancer and starting age of screening based on lifestyle, environmental, and genetic factors. Gastroenterology 2018;154:2152–2164.e19.
21. Weigl K, Thomsen H, Balavarca Y, et al. Genetic risk ccore is associated with prevalence of advanced neoplasms in a colorectal cancer screening population. Gastroenterology 2018;155:88–98 e10.
22. Guo F, Edelmann D, Cardoso R, et al. Polygenic risk score for defining personalized surveillance intervals after adenoma detection and removal at colonoscopy. Clin Gastroenterol Hepatol 2023;21:210–219.e11.
23. Guo F, Weigl K, Carr PR, et al. Use of polygenic risk scores to select screening intervals after negative findings from colonoscopy. Clin Gastroenterol Hepatol 2020; 18:2742–2751.e7.
24. Spira A, Yurgelun MB, Alexandrov L, et al. Precancer atlas to drive precision prevention trials. Cancer Res 2017;77:1510–1541.
25. Lieberman DA, Weiss DG, Bond JH, et al. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. N Engl J Med 2000;343:162–168.
26. Demb J, Gupta S. Realizing the promise of personalized colorectal cancer screening in practice. J Natl Cancer Inst 2021;113:1120–1122.
27. Sullivan BA, Qin X, Redding TS, et al. Genetic colorectal cancer and adenoma risk variants are associated with increasing adenoma counts. Cancer Epidemiol Biomarkers Prev 2020;29:2269–2276.
28. Thomas M, Sakoda LC, Hoffmeister M, et al. Genome-wide modeling of polygenic risk score in colorectal cancer risk. Am J Hum Genet 2020;107:432–444.
29. Jiao S, Hsu L, Hutter CM, et al. The use of imputed values in the meta-analysis of genome-wide association studies. Genet Epidemiol 2011;35:597–605.
30. Taliun D, Harris DN, Kessler MD, et al. Sequencing of 53, 831 diverse genomes from the NHLBI TOPMed Program. Nature 2021;590:290–299.
31. Hunter-Zinck H, Shi Y, Li M, et al. Genotyping array design and data quality control in the Million Veteran Program. Am J Hum Genet 2020;106:535–548.
32. Saunders CL, Kilian B, Thompson DJ, et al. External validation of risk prediction models incorporating common genetic variants for incident colorectal cancer using UK Biobank. Cancer Prev Res 2020; 13:509–520.
33. Lee JK, Liles EG, Bent S, et al. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Intern Med 2014; 160:171.

34. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. N Eng J Med 2012;366:697–706.
35. Makaroff KE, Shergill J, Lauzon M, et al. Patient preferences for colorectal cancer screening tests in light of lowering the screening age to 45 years. Clin Gastroenterol Hepatol 2023;21:520–531.e10.
36. Chen F, Darst BF, Madduri RK, et al. Validation of a multi-ancestry polygenic risk score and age-specific risks of prostate cancer: a meta-analysis within diverse populations. Elife 2022;11:e78304.
37. Conti DV, Darst BF, Moss LC, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. Nat Genet 2021;53:65–75.
38. Wang L, Desai H, Verma SS, et al. Performance of polygenic risk scores for cancer prediction in a racially diverse academic biobank. Genet Med 2022;24:601–609.
39. Mur P, Bonifaci N, Díez-Villanueva A, et al. Non-Lynch familial and early-onset colorectal cancer explained by accumulation of low-risk genetic variants. Cancers (Basel) 2021;13:3857.
40. McGeoch L, Saunders CL, Griffin SJ, et al. Risk prediction models for colorectal cancer incorporating common genetic variants: a systematic review. Cancer Epidemiol Biomarkers Prev 2019;28:1580–1593.
41. Gupta S, Thrift AP. Polygenic risk scores for follow up after colonoscopy and polypectomy: another tool for risk stratification and planning surveillance? Clin Gastroenterol Hepatol 2023;21:29–32.
42. Gupta S, Earles A, Bustamante R, et al. Adenoma detection rate and clinical characteristics influence advanced neoplasia risk after colorectal polypectomy. Clin Gastroenterol Hepatol 2023;21:1924–1936.e9.
43. Dixon P, Keeney E, Taylor JC, et al. Can polygenic risk scores contribute to cost-effective cancer screening? A systematic review. Genet Med 2022;24:1604–1617.

Received September 21, 2023. Accepted October 3, 2023.

Correspondence:
Address correspondence to: Brian Sullivan, MD, MHS, Duke University School of Medicine, 2301 Erwin Rd, Durham, North Carolina 27710. e-mail: brian.a.sullivan@duke.edu.

Acknowledgments:
The authors would like to acknowledge and thank the participants of CSP#380 for their generosity in contributing to this ongoing, longitudinal cohort study. This work is dedicated to Dawn Provenzale, MD, MS, former Principal Investigator of CSP#380, for her dedication and contributions to cancer research. The views presented in this manuscript are those of the authors and do not necessarily represent the position or policy of the Department of Veterans Affairs or the United States Government.

Authors' Contributions:
Study concept and design: Brian A. Sullivan, Xuejun Qin, David A. Lieberman, and Elizabeth R. Hauser. Acquisition of data: Brian A. Sullivan, Xuejun Qin, Thomas S. Redding IV, David Weiss, Julie Upchurch, Kellie J. Sims, Anjanette Stone, Belinda Ear, Christina D. Williams, David A. Lieberman, and Elizabeth R. Hauser. Analysis and interpretation of data: Brian A. Sullivan, Xuejun Qin, Thomas S. Redding IV, David Weiss, Jason A. Dominitz, Anjanette Stone, David A. Lieberman, and Elizabeth R. Hauser. Drafting of manuscript: Brian A. Sullivan, David Weiss, Julie Upchurch, Jason A. Dominitz, Belinda Ear, and Elizabeth R. Hauser. Critical revision of the manuscript for important intellectual content: Xuejun Qin, Thomas S. Redding IV, David Weiss, Kellie J. Sims, Jason A. Dominitz, Christina D. Williams, David A. Lieberman, and Elizabeth R. Hauser. Statistical analysis: Xuejun Qin, Thomas S. Redding IV, and Elizabeth R. Hauser. Obtained funding: Anjanette Stone, Christina D. Williams, David A. Lieberman, and Elizabeth R. Hauser. Administrative, technical, or material support: Julie Upchurch, Kellie J. Sims, and Belinda Ear. Study supervision: Jason A. Dominitz, Christina D. Williams, David A. Lieberman, and Elizabeth R. Hauser.

Conflicts of Interest:
These authors disclose the following: Brian A. Sullivan reports grant support from Exact Sciences, which is outside the submitted work. David A. Lieberman reports other support from Check-Cap, Ironwood, ColoWrap, and Freenome which is outside the submitted work. The remaining authors disclose no conflicts.

Funding:
Brian A. Sullivan is supported by the AGA Research Foundation's AGA Research Scholar Award–AGA2021-13-03. This work was also funded by the US Department of Veteran Affairs Cooperative Studies Program.

Ethical Statement:
The corresponding author, on behalf of all authors, jointly and severally, certifies that their institution has approved the protocol for any investigation involving humans or animals and that all experimentation was conducted in conformity with ethical and humane principles of research. Brian Sullivan, MD, MHS, is the guarantor of the article. This protocol is approved by the Durham VA Institutional Review Board under the CSP#380b study (#01797/0002).

Data Transparency Statement:
Analysis code is available upon reasonable request (https://www.vacsp.research.va.gov/CSPEC/Studies/INVESTD-R/CSP-380-Risk-Factor-Colonic-Adenomas.asp).

Disclaimer:
The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or the government of the United States.

Reporting Guidelines:
STROBE.